

WATERMAN

(TK)

Release 3.1A John F. Collins, BioComputing Research Unit,
Copyright (C) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPatch_app protein - protein database search, using Smith-Waterman algorithm
Run On: Mon Jun 19 16:01:00 2000 MasPar time 3.40 seconds
215.791 Million cell updates/sec
Inuclear output not generated.

Title: >US-09-142-524A-4
Description: (1-31); from USC9142524A.pep
Perfect Score: 225
Sequence: 1 IFSKNLNKLNMPLYIAGNKRRIKRVSNVI 31

Scoring table: FAM 150
Gap 11

Searched: 189963 seqs, 23696106 residues

Post-processing: Minimum Match: 0.1
Listing first 45 summaries

Database: a-geneseq35
l-geneseqf

Statistics: Mean 23.433; Variance 89.397; scale 0.252

Prof. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description	Prof. No.
1	225	100.0	31	W27372	Multi-epitope peptide	4.18e-04
2	143	63.6	20	W42130	T-cell epitope peptide	9.27e-06
3	143	63.6	354	W04344	Chamaecyparis obtusa p	9.27e-06
4	143	63.6	354	W42121	Japanese cypress pollen	9.27e-06
5	143	63.6	375	W04345	Chamaecyparis obtusa p	9.27e-06
6	123	54.7	367	R45577	Jun s I.	8.28e-04
7	123	54.7	370	R45578	Jun v I.	8.28e-04
8	116	52.4	134	W27371	Multi-epitope peptide	2.50e-03
9	116	52.4	353	R75366	Japanese cedar pollen	2.50e-03
10	116	52.4	353	R41387	Cedar pollen allergen	2.50e-03
11	116	52.4	374	R31937	Cry j I.	2.50e-03
12	116	52.4	374	R60466	Japanese cedar pollen	2.50e-03
13	116	52.4	374	R82490	Cry j I Japanese Cedar	2.50e-03
14	116	52.4	374	R45541	Cry j I pollen allergen	2.50e-03
15	113	50.2	20	R45549	Cry j I pollen allergen	2.48e-03
16	113	50.2	20	R82498	Cry j I Japanese cedar	2.48e-03
17	113	50.2	20	R45542	T-cell epitope peptide	2.48e-03
18	87	38.7	105	W27370	Multi-epitope peptide	1.89e-00
19	92	36.4	31	W27373	Multi-epitope peptide	5.23e-00
20	92	35.6	80	W27369	Multi-epitope peptide	7.84e-00
21	79	35.1	1018	W06485	Rat contactin ligand f	9.59e-01
22	74	32.9	20	W42131	T-cell epitope peptide	2.59e-01
23	74	32.9	637	W52294	Drosophila p7356K prot	2.59e-01

ALIGNMENTS

RESULT 1
ID W27372 standard; peptide: 31 AA
AC W27372: 1 238 32.4 73 32.4 238 1 4755
25 73 32.4 238 1 4755
26 72 32.0 963 1 4755
27 71 31.6 615 1 4755
28 70 31.1 513 1 4755
29 70 31.1 513 1 4755
30 70 31.1 721 1 4755
31 70 31.1 746 1 4755
32 69 30.7 20 1 4755
33 69 30.7 635 1 4755
34 69 30.7 635 1 4755
35 69 30.7 635 1 4755
36 69 30.7 635 1 4755
37 69 30.7 1029 1 4755
38 68 30.2 509 1 4755
39 68 30.2 509 1 4755
40 68 30.2 1018 1 4755
41 68 30.2 1018 1 4755
42 67 29.8 738 1 4755
43 67 29.8 738 1 4755
44 67 29.8 738 1 4755
45 67 29.8 738 1 4755
Human thoracic aorta G 3.15e-01
Human thoracic aorta G 3.15e-01
Thrombostable enzyme (m 3.15e-01
Tomato IgE1 ethylene r 3.15e-01
QETR ethylene response 3.15e-01
A. thaliana ethylene r 3.15e-01
Racilius subtilis telic 3.15e-01
Racilius subtilis telic 3.15e-01
T-cell epitope peptide 3.15e-01
Tomato ethylene respon 3.15e-01
Tomato ethylene respon 3.15e-01
Nr (cavear ripe) tomato 3.15e-01
Tomato ethylene respon 3.15e-01
Homo sapiens D18511 c 3.15e-01
Human SR-B1 class B sc 3.15e-01
Human SR-B1 class B sc 3.15e-01
Human contactin (EXHL 3.15e-01
Human contactin. 3.15e-01
Ethylene response (ETR 3.15e-01
Ethylene response (ETR 3.15e-01
Ethylene response (ETR 3.15e-01
A. thaliana ethylene r 3.15e-01

Query Match 100.0%; Score 225; DB 1; Length 31;
Best Local Similarity 100.0%; Pred. No. 4.18e-04;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 1 IFSKNLNKLNMPLYIAGNRRRIKRVSNVI 31
Qy 1 IFSKNLNKLNMPYIAGNRRRIKRVSNVI 31
RESULT 2
ID W42130 standard; peptide: 20 AA
AC W42130: 1 238 32.4 73 32.4 238 1 4755
25 73 32.4 238 1 4755
26 72 32.0 963 1 4755
27 71 31.6 615 1 4755
28 70 31.1 513 1 4755
29 70 31.1 513 1 4755
30 70 31.1 721 1 4755
31 70 31.1 746 1 4755
32 69 30.7 20 1 4755
33 69 30.7 635 1 4755
34 69 30.7 635 1 4755
35 69 30.7 635 1 4755
36 69 30.7 635 1 4755
37 69 30.7 1029 1 4755
38 68 30.2 509 1 4755
39 68 30.2 509 1 4755
40 68 30.2 1018 1 4755
41 68 30.2 1018 1 4755
42 67 29.8 738 1 4755
43 67 29.8 738 1 4755
44 67 29.8 738 1 4755
45 67 29.8 738 1 4755
Human thoracic aorta G 3.15e-01
Human thoracic aorta G 3.15e-01
Thrombostable enzyme (m 3.15e-01
Tomato IgE1 ethylene r 3.15e-01
QETR ethylene response 3.15e-01
A. thaliana ethylene r 3.15e-01
Racilius subtilis telic 3.15e-01
Racilius subtilis telic 3.15e-01
T-cell epitope peptide 3.15e-01
Tomato ethylene respon 3.15e-01
Tomato ethylene respon 3.15e-01
Nr (cavear ripe) tomato 3.15e-01
Tomato ethylene respon 3.15e-01
Homo sapiens D18511 c 3.15e-01
Human SR-B1 class B sc 3.15e-01
Human SR-B1 class B sc 3.15e-01
Human contactin (EXHL 3.15e-01
Human contactin. 3.15e-01
Ethylene response (ETR 3.15e-01
Ethylene response (ETR 3.15e-01
Ethylene response (ETR 3.15e-01
A. thaliana ethylene r 3.15e-01

KW Japanese cedar; pollen allergen; allergy treatment; diagnosis;
KW T cell epitope; sensitivity; detection;
OS Juniperus sabinoides.
FH Key Location/Qualifiers
FT peptide 1..21 /note="signal peptide"
FT peptide 22..367 /note="mature peptide"
FN W09401560-A.
PD 20-JAN-1994.
PF 15-JAN-1993: U00139.
PR 10-JUL-1992: W0-J05661.
PA (IMMU-) IMMUNOLOGIC PHARM CORP.
PI Bond JF, Gorman RD, Griffith LJ, Kuo M, Pollock J.
DR WPI: 94-035065/04.
DR N-PSDB: Q55272.
PI Antigens derived from Japanese cedar pollen allergen Cry j 1 -
PI contain at least two T cell epitope(s), used to treat or diagnose
PI allergy.
PS Disclosure: Fig 16: 137pp; English.
CC The sequence is that of Jun s 1, a homologue of the Japanese
CC cedar pollen allergen Cry j 1. Antigenic peptides derived from it
CC can be used for the treatment and diagnosis of allergies associated
CC with Japanese cedar pollen.
SQ Sequence 367 AA:

Query Match 54.7% Score 123 DB 1 Length 367

Best Local Similarity 75.0% Pred. No. 8.28e-04

Matches 15: Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 92 IFSNNMNIKMPYVAGHK 11

11111111111111111111

QY 1 IFSKNLNKLNMPLYIAGNK 20

RESULT 7
ID R45578 standard; Protein: 370 AA.

AC R45578

DT 23-JUL-1994 (first entry)

DE Jun v 1

KW Japanese cedar; pollen allergen; allergy treatment; diagnosis;

KW T cell epitope; sensitivity; detection;

OS Juniperus virginiana.

FH Key Location/Qualifiers

FT peptide 1..21

FT peptide /note="signal peptide"

FT peptide 22..370

FT peptide /note="mature peptide"

FN W09401560-A.

PD 20-JAN-1994.

PF 15-JAN-1993: U00139

PR 10-JUL-1992: W0-J05661.

PA (IMMU-) IMMUNOLOGIC PHARM CORP.

PI Bond JF, Gorman RD, Griffith LJ, Kuo M, Pollock J.

DR WPI: 94-035065/04.

DR N-PSDB: Q55273.

PI Antigens derived from Japanese cedar pollen allergen Cry j 1 -

PI contain at least two T cell epitope(s), used to treat or diagnose

PI allergy

PS Disclosure: Fig 17: 137pp; English.

CC The sequence is that of Jun v 1, a homologue of the Japanese

CC cedar pollen allergen Cry j 1. Antigenic peptides derived from it

CC can be used for the treatment and diagnosis of allergies associated

CC with Japanese cedar pollen.

SQ Sequence 370 AA:

Query Match 54.7% Score 123 DB 1 Length 370

Best Local Similarity 75.0% Pred. No. 8.28e-04

Matches 15: Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 92 IFSNNMNIKMPYVAGHK 11

11111111111111111111

QY 1 IFSKNLNKLNMPLYIAGNK 20

QY 1 IFSKNLNKLNMPLYIAGNK 20

RESULT 8

ID W27371 standard; peptide: 134 AA.

AC W27371

DT 24-MAR-1998 (first entry)

DE Multi-epitope peptide used as immunotherapeutic agent #3

KW Multi-epitope peptide; immunotherapeutic agent; allergic disease;

KW T-cell epitope region; allergen; lymphocyte; immunoglobulin E;

OS Synthetic.

PN W09732600-A1.

PD 12-SEP-1997.

PF 10-MAR-1997: J00740.

PR 10-MAR-1996: JP-080702.

PA (MEIP) MEIJI MILK PROD CO LTD.

PI Daiiki K, Iwama A, Kuro K, Kure A, Sone T;

DR WPI: 97-470495/43.

PI Peptide immunotherapeutic agent to treat allergic diseases -

PI contains multi-epitope peptide containing T cell epitope regions

PS Claim 6: Page 32: 58pp; Japanese.

CC The present sequence represents a multi-epitope peptide which is used as

CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2

CC or more different allergens (preferably linked via arginine or lysine

CC dimers), where the T cell epitope regions have a positivity index

CC greater than 100 as measured in a patient group responding to the

CC allergen; have at least 70% reactivity with lymphocytes from patients

CC responding to the allergen; and are not reactive with immunoglobulin E

CC (IgE) antibodies from patients responsive to the allergen. The agent can

CC be used to prevent and treat a wide variety of allergic diseases, e.g. by

CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.

SQ Sequence 134 AA:

Query Match 52.4% Score 118 DB 1 Length 134

Best Local Similarity 60.0% Pred. No. 2.50e-03

Matches 19: Conservative 9; Mismatches 6; Indels 1; Gaps 1;

Db 78 IFSNNMNIKMPYIAGYKT-PDGHRAEV 106

11111111111111111111

QY 1 IFSKNLNKLNMPLYIAGNKRRIKVSNV 30

RESULT 9

ID R45578 standard; protein: 453 AA.

AC R45578

DT 12-MAR-1996 (first entry)

DE Japanese cedar pollen allergen Cry j 1

KW Japanese cedar; pollen allergen; Cry j 1; T-cell epitope; peptides;

KW prevention; treatment; cryptomeria pollinosis.

OS Cryptomeria japonica.

FH Key Location/Qualifiers

FT peptide 61..75

FT peptide /note="T-cell epitope peptide"

FT peptide 91..105

FT peptide /note="T-cell epitope peptide"

FT peptide 106..120

FT peptide /note="T-cell epitope peptide"

FT peptide 146..160

FT peptide /note="T-cell epitope peptide"

FT peptide 211..225

FT peptide /note="T-cell epitope peptide"

FT peptide 326..340

FT peptide /note="T-cell epitope peptide"

FT peptide 335..346

FT peptide /note="T-cell epitope peptide"

PN J0718295-A.

PD 09-MAY-1995.

PF 20-OCT-1993: 262526.

PR 20-OCT-1993: JP-262626.

PA (MEIP) MEIJI MILK PROD CO LTD.

DR WPI: 95-203834/27.

PT New cryptomeria pollen allergen T-cell epitope peptide - used for prevention, treatment and investigation of Japanese cedar pollinosis
 PS Disclature: Kios 1-2: 8pp; Japanese
 CC R7588 is the Japanese cedar pollen allergen Cryj 1, from which the T-cell epitope peptides R89283-R89295 were derived. The peptides can be used for the prevention and treatment of cryptomeria pollinosis, and also for the investigation of pollinosis.
 SQ Sequence 353 AA;

Query Match 52.4%; Score 118; DB 1; Length 353;

Best Local Similarity 60.0%; Pred. NO. 2,50e-03;

Matches 18; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

DB 71 IFSGNMNIKLMKMPYIAGYKT-FDGRGACV 99

QY 1 IFSKNLNKLNMPYIAGNKRRIKRVSNV 30

RESULT 10

ID R81587 standard: Protein: 353 AA.

AC R81587

DT 24-MAY-1996 (first entry)

DE Cedar pollen allergen B.

KW Cedar pollen allergen; immunoglobulin E; IgE; T-cell epitope;

OS Antibody; pollinosis; therapy; immunotherapy.

CS Cryptomeria japonica.

PN EP-700929-A2.

PD 13-MAY-1996.

DE 08-SEP-1995; 306295.

DE 10-SEP-1994; JP-242137.

DE 14-JUL-1995; JP-200221.

DE 14-JUL-1995; JP-200204.

PA (HAYB) : HAYASHI; RARA; SEIBUITSU KAGAKU.

PI Hino K., Saito S., Taniguchi Y.

DB 74-14078/15.

DT New peptide(s) derived from cedar pollen allergens - activate allergen-specific T-cells, but not allergen-specific IgE antibodies, used for treating cedar pollinosis.

PS Claim 5; Page 31-32; 36pp; English.

CC Synthetic peptides based on portions of cedar pollen allergens A (R81585) and B (R81587) were tested for their ability to activate cedar allergen-specific T-cells, but not allergen-specific IgE antibodies. 6 peptides (R81586-R81589) were identified as T-cell epitopes. These peptides, plus subsequences (R81573-79) essential for T-cell recognition, and homologous peptides (R81588-96) can be used as immunotherapeutic agents to treat or prevent cedar pollinosis, avoiding side-effects such as anaphylaxis.

SQ Sequence 353 AA;

Query Match 52.4%; Score 118; DB 1; Length 353;

Best Local Similarity 60.0%; Pred. NO. 2,50e-03;

Matches 18; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

DB 71 IFSGNMNIKLMKMPYIAGYKT-FDGRGACV 99

QY 1 IFSKNLNKLNMPYIAGNKRRIKRVSNV 30

RESULT 11

ID R81937 standard: Protein: 374 AA.

AC R81937

DT 03-JUN-1993 (first entry)

DE Cry j 1.

KW Japanese cedar pollen allergen; antigen; allergy; B cell; T cell.

OS Cryptomeria japonica.

PH Key Location/Qualifiers

FI Peptide 1, 21

FI protein /note="signal peptide"

FI protein 22, 374

FI protein /note="mature Cry j 1"

PN W093012.3-A

PD 21-JAN-1993.

PT 10-JUL-1992; J05661.

PR 12-JUL-1991; US-729134.
 PR 15-JUL-1991; US-730452.
 PA (IMMO-) IMMUNOLOGIC PHARM CORP.
 PI Bond JF., Griffith J.J., Pollock J.
 DR WPI: 93-045434/05.
 DR N-PSDB: Q35304.
 PI Nucleic acid sequence encoding cryptomeria japonica allergen - for the diagnosis treatment and prevention of allergic reactions to Japanese cedar pollen.
 PT Claim 11; Page 42; 69pp; English.
 CC Fresh pollen and staminate cone samples were collected from a single Cryptomeria japonica (Japanese cedar) tree. RNA was prep'd and used to synthesize cDNA. The cDNA was subjected to successive rounds of PCR to yield a full length Cry j 1 clone. Cry j 1 or an antigenic fragment of it may be used for detecting, treating and preventing an allergic response to Japanese cedar pollen allergen. It is capable of modifying both the B and T cell response to Cry j 1 and T cell response to a Cry j 1 antigen.
 SQ Sequence 374 AA;

Query Match 52.4%; Score 118; DB 1; Length 374;

Best Local Similarity 60.0%; Pred. NO. 2,50e-03;

Matches 18; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

DB 92 IFSGNMNIKLMKMPYIAGYKT-FDGRGACV 120

QY 1 IFSKNLNKLNMPYIAGNKRRIKRVSNV 30

RESULT 12

ID R60166 standard: Protein: 374 AA.

AC R60166

DT 24-MAR-1995 (first entry)

DE Japanese cedar pollen antigen Cryj1.

KW Japanese cedar pollen antigen; allergen; Cryj1; suai; pollinosis.

OS Cryptomeria japonica.

PH Key Location/Qualifiers

FI Peptide 1, 21

FI protein /label="signal peptide"

FI protein 22, 374

FI protein /label="mature Cryj1"

PN J06197768-A.

PD 19-JUL-1994.

PR 07-JAN-1993; J01116.

PR 07-JAN-1993; JP-C01116.

PA (MEIJ) MEIJ; SEIKA KAISHA.

DR WPI: 94-268680/34.

DR N-PSDB: Q71601.

PI Suai (Japanese cedar) pollen antigen Cryj1 is useful in

diagnosis, treatment and prevention of suai pollinosis.

PS Claim 2; Page 5-7; 9pp; Japanese.

CC The coding sequence for the Japanese cedar ("suai") pollen allergen

Cryj1 was isolated from a cDNA library prepared from PO-YA mRNA. A

or part of the Cryj1 protein can be used for diagnosis, treatment

and prevention of suai pollinosis.

SQ Sequence 374 AA;

Query Match

Best Local Similarity 60.0%; Pred. NO. 2,50e-03;

Matches 18; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

DB 92 IFSGNMNIKLMKMPYIAGYKT-FDGRGACV 120

QY 1 IFSKNLNKLNMPYIAGNKRRIKRVSNV 30

RESULT 13

ID R82490 standard: Protein: 374 AA.

AC R82490

DT 15-APR-1996 (first entry)

DE Cry j 1 Japanese Cedar pollen allergen.

KW Cry j 1; Japanese cedar pollen allergen; modified; drug production;

allergy; Cryptomeria japonica.

Search completed: Mon Jun 19 16:01:15 2000
Job time : 15 secs.

1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 20. 21. 22. 23. 24. 25. 26. 27. 28. 29. 30. 31. 32. 33. 34. 35. 36. 37. 38. 39. 40. 41. 42. 43. 44. 45. 46. 47. 48. 49. 50. 51. 52. 53. 54. 55. 56. 57. 58. 59. 60. 61. 62. 63. 64. 65. 66. 67. 68. 69. 70. 71. 72. 73. 74. 75. 76. 77. 78. 79. 80. 81. 82. 83. 84. 85. 86. 87. 88. 89. 90. 91. 92. 93. 94. 95. 96. 97. 98. 99. 100. 101. 102. 103. 104. 105. 106. 107. 108. 109. 110. 111. 112. 113. 114. 115. 116. 117. 118. 119. 120. 121. 122. 123. 124. 125. 126. 127. 128. 129. 130. 131. 132. 133. 134. 135. 136. 137. 138. 139. 140. 141. 142. 143. 144. 145. 146. 147. 148. 149. 150. 151. 152. 153. 154. 155. 156. 157. 158. 159. 160. 161. 162. 163. 164. 165. 166. 167. 168. 169. 170. 171. 172. 173. 174. 175. 176. 177. 178. 179. 180. 181. 182. 183. 184. 185. 186. 187. 188. 189. 190. 191. 192. 193. 194. 195. 196. 197. 198. 199. 200. 201. 202. 203. 204. 205. 206. 207. 208. 209. 210. 211. 212. 213. 214. 215. 216. 217. 218. 219. 220. 221. 222. 223. 224. 225. 226. 227. 228. 229. 230. 231. 232. 233. 234. 235. 236. 237. 238. 239. 240. 241. 242. 243. 244. 245. 246. 247. 248. 249. 250. 251. 252. 253. 254. 255. 256. 257. 258. 259. 260. 261. 262. 263. 264. 265. 266. 267. 268. 269. 270. 271. 272. 273. 274. 275. 276. 277. 278. 279. 280. 281. 282. 283. 284. 285. 286. 287. 288. 289. 290. 291. 292. 293. 294. 295. 296. 297. 298. 299. 300. 301. 302. 303. 304. 305. 306. 307. 308. 309. 310. 311. 312. 313. 314. 315. 316. 317. 318. 319. 320. 321. 322. 323. 324. 325. 326. 327. 328. 329. 330. 331. 332. 333. 334. 335. 336. 337. 338. 339. 340. 341. 342. 343. 344. 345. 346. 347. 348. 349. 350. 351. 352. 353. 354. 355. 356. 357. 358. 359. 360. 361. 362. 363. 364. 365. 366. 367. 368. 369. 370. 371. 372. 373. 374. 375. 376. 377. 378. 379. 380. 381. 382. 383. 384. 385. 386. 387. 388. 389. 390. 391. 392. 393. 394. 395. 396. 397. 398. 399. 400. 401. 402. 403. 404. 405. 406. 407. 408. 409. 410. 411. 412. 413. 414. 415. 416. 417. 418. 419. 420. 421. 422. 423. 424. 425. 426. 427. 428. 429. 430. 431. 432. 433. 434. 435. 436. 437. 438. 439. 440. 441. 442. 443. 444. 445. 446. 447. 448. 449. 450. 451. 452. 453. 454. 455. 456. 457. 458. 459. 460. 461. 462. 463. 464. 465. 466. 467. 468. 469. 470. 471. 472. 473. 474. 475. 476. 477. 478. 479. 480. 481. 482. 483. 484. 485. 486. 487. 488. 489. 490. 491. 492. 493. 494. 495. 496. 497. 498. 499. 500. 501. 502. 503. 504. 505. 506. 507. 508. 509. 510. 511. 512. 513. 514. 515. 516. 517. 518. 519. 520. 521. 522. 523. 524. 525. 526. 527. 528. 529. 530. 531. 532. 533. 534. 535. 536. 537. 538. 539. 540. 541. 542. 543. 544. 545. 546. 547. 548. 549. 550. 551. 552. 553. 554. 555. 556. 557. 558. 559. 560. 561. 562. 563. 564. 565. 566. 567. 568. 569. 570. 571. 572. 573. 574. 575. 576. 577. 578. 579. 580. 581. 582. 583. 584. 585. 586. 587. 588. 589. 590. 591. 592. 593. 594. 595. 596. 597. 598. 599. 600. 601. 602. 603. 604. 605. 606. 607. 608. 609. 610. 611. 612. 613. 614. 615. 616. 617. 618. 619. 620. 621. 622. 623. 624. 625. 626. 627. 628. 629. 630. 631. 632. 633. 634. 635. 636. 637. 638. 639. 640. 641. 642. 643. 644. 645. 646. 647. 648. 649. 650. 651. 652. 653. 654. 655. 656. 657. 658. 659. 660. 661. 662. 663. 664. 665. 666. 667. 668. 669. 670. 671. 672. 673. 674. 675. 676. 677. 678. 679. 680. 681. 682. 683. 684. 685. 686. 687. 688. 689. 690. 691. 692. 693. 694. 695. 696. 697. 698. 699. 700. 701. 702. 703. 704. 705. 706. 707. 708. 709. 710. 711. 712. 713. 714. 715. 716. 717. 718. 719. 720. 721. 722. 723. 724. 725. 726. 727. 728. 729. 730. 731. 732. 733. 734. 735. 736. 737. 738. 739. 740. 741. 742. 743. 744. 745. 746. 747. 748. 749. 750. 751. 752. 753. 754. 755. 756. 757. 758. 759. 760. 761. 762. 763. 764. 765. 766. 767. 768. 769. 770. 771. 772. 773. 774. 775. 776. 777. 778. 779. 780. 781. 782. 783. 784. 785. 786. 787. 788. 789. 790. 791. 792. 793. 794. 795. 796. 797. 798. 799. 800. 801. 802. 803. 804. 805. 806. 807. 808. 809. 810. 811. 812. 813. 814. 815. 816. 817. 818. 819. 820. 821. 822. 823. 824. 825. 826. 827. 828. 829. 830. 831. 832. 833. 834. 835. 836. 837. 838. 839. 840.

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

```

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Tue Jun 20 13:34:00 2000: MasPar time 2.66 Seconds
168.026 Million cell updates/sec
Tabular output not generated.

```

Title: US-09-142-524A-4
Description: (1-3) from US09142524A.pep
Perfect Score: 225
sequence: 1 IESKLNLIKLNMPYIAGNKRREIKRVSNI 31

Scoring table: PAM 150
Gap 11

Searched: 14534; seqs, 14437480 residues

| Post-processing | Minimum Match | Match | Summary |
|-----------------|---------------|---------|---------|
| 08 | 08 | 08 | 08 |
| 15 | 15 | 15 | 15 |
| 45 | 45 | 45 | 45 |
| Summary | Summary | Summary | Summary |

```
Database:
4-issued
1:5A COMB 2:5B COMB 3:5 COMB 4:PCT COMB 5:backfiles
```

Statistics: Mean 21.952; Variance 87.801; scale 0.253

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | | Length | DB | ID | Description | Freq. No. | |
|------------|-------|----|--------|------|-------------|-------------|-------------|----------|
| | Match | Q | | | | | | |
| 1 | 79 | 35 | 1058 | 1 | US-08-4152- | Sequence 2 | 6.12e-00 | |
| 2 | 73 | 32 | 298 | 1 | US-08-118- | Sequence 76 | 1.97e-01 | |
| 3 | 73 | 32 | 4 | 298 | 4 | PCR-US93-9 | Sequence 76 | 1.97e-01 |
| 4 | 72 | 32 | 963 | 1 | US-08-537- | Sequence 3 | 2.38e-01 | |
| 5 | 70 | 31 | 613 | 2 | US-08-434- | Sequence 48 | 3.49e-01 | |
| 6 | 70 | 31 | 613 | 2 | US-08-484- | Sequence 45 | 3.49e-01 | |
| 7 | 69 | 30 | 725 | 2 | US-08-749- | Sequence 7 | 4.22e-01 | |
| 8 | 69 | 30 | 725 | 2 | US-08-749- | Sequence 8 | 4.22e-01 | |
| 9 | 69 | 30 | 735 | 2 | US-08-484- | Sequence 35 | 4.22e-01 | |
| 10 | 69 | 30 | 735 | 2 | US-08-484- | Sequence 50 | 4.22e-01 | |
| 11 | 68 | 30 | 2 | 509 | 2 | US-08-890- | Sequence 2 | 5.09e-01 |
| 12 | 68 | 30 | 2 | 605 | 2 | US-08-732- | Sequence 8 | 5.09e-01 |
| 13 | 68 | 30 | 2 | 1018 | 3 | US-08-040- | Sequence 6 | 5.09e-01 |
| 14 | 68 | 30 | 2 | 1018 | 1 | US-08-714- | Sequence 6 | 5.09e-01 |
| 15 | 68 | 30 | 2 | 1018 | 1 | US-08-408- | Sequence 6 | 5.09e-01 |
| 16 | 68 | 30 | 2 | 1018 | 1 | US-08-408- | Sequence 6 | 5.09e-01 |
| 17 | 67 | 29 | 8 | 155 | 1 | US-08-530- | Sequence 12 | 5.15e-01 |
| 18 | 67 | 29 | 8 | 155 | 2 | US-08-484- | Sequence 12 | 5.15e-01 |
| 19 | 67 | 29 | 8 | 738 | 2 | US-08-484- | Sequence 7 | 6.15e-01 |
| 20 | 67 | 29 | 8 | 738 | 2 | US-08-484- | Sequence 9 | 6.15e-01 |
| 21 | 67 | 29 | 8 | 738 | 1 | US-08-530- | Sequence 9 | 6.15e-01 |
| 22 | 67 | 29 | 8 | 738 | 2 | US-08-484- | Sequence 3 | 6.15e-01 |
| 23 | 67 | 29 | 8 | 738 | 1 | US-08-530- | Sequence 7 | 6.15e-01 |

S2 SEQUENCE 113102 MW: 5593224 CN:
 Query Match 35.1% Score 79: DB 1: Length 1018:
 Best Local Similarity 34.5% Pred. No. 1: 97e-01:
 Matches 10: Conservative 11: Mismatches 6: Indels 2: Gaps 2:

Db 164 PPSLSYKALNNGTFLGSPDYVRRVSRIV 192
 QY 2 PPSKLNLIKLNWPLYAGNKKRFKRVSNVI 28

RESULT 2
 ID US-08-118-276 STANDARD: PRT: 298 AA:
 XX XXXXXX
 DE Sequence 76, Application US/08118270
 CC Patent No. 5508344
 CC GENERAL INFORMATION:
 CC APPLICANT: Murphy, Randall B.
 CC APPLICANT: Schuster, David I.
 CC TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
 CC TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
 CC NUMBER OF SEQUENCES: 348
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: BROWDY AND NEIMARK
 CC STREET: 419 Seventh Street, N.W., Suite 300
 CC CITY: Washington
 CC STATE: D.C.
 CC COUNTRY: USA
 CC ZIP: 20004
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patent In Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/118 270
 CC FILING DATE: 09-SEP-1993
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: US 07/943,246
 CC FILING DATE: 10-SEP-1992
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Townsend, Kevin G.
 CC REGISTRATION NUMBER: 34,033
 CC REFERENCE/DOCKET NUMBER: MURPHY-2A
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: 202-628-5197
 CC TELEFAX: 202-737-3528
 CC TELEX: 248633
 CC INFORMATION FOR SEQ ID NO: 76:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 298 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS: single
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: peptide
 CC SEQUENCE 298 AA: 33548 MW: 467538 CN:

Query Match 32.4% Score 73: DB 1: Length 298:
 Best Local Similarity 32.3% Pred. No. 1: 97e-01:
 Matches 10: Conservative 10: Mismatches 10: Indels 1: Gaps 1:

Db 50 LFSKAVIALNNGTFLGSPDYVRRVSRIV 79
 QY 1 LFSKLNLIKLNWPLYAGNKKRFKRVSNVI 31

RESULT 4
 ID US-08-537-002A-3 STANDARD: PRT: 963 AA:
 XX XXXXXX
 DE Sequence 3, Application US/08537002A
 CC Sequence 3, Application US/08537002A
 CC Patent No. 5773282
 CC GENERAL INFORMATION:
 CC APPLICANT: TSUSAKI, Keiji
 CC APPLICANT: KUBOTA, Michio
 CC APPLICANT: SUGIMOTO, Toshiyuki

ID PCT-US93-08528-76 STANDARD: PRT: 298 AA:
 XX XXXXXX
 DE Sequence 76, Application PCT/US93/08528
 CC Sequence 76, Application PCT/US93/08528
 CC GENERAL INFORMATION:
 CC APPLICANT: New York University
 CC TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
 CC TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
 CC NUMBER OF SEQUENCES: 348
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: BROWDY AND NEIMARK
 CC STREET: 419 Seventh Street, N.W., Suite 300
 CC CITY: Washington
 CC STATE: D.C.
 CC COUNTRY: USA
 CC ZIP: 20004
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patent In Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: PCT/US93/08528
 CC FILING DATE: 09-SEP-1993
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: US 07/943,246
 CC FILING DATE: 10-SEP-1992
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Townsend, Kevin G.
 CC REGISTRATION NUMBER: 34,033
 CC REFERENCE/DOCKET NUMBER: MURPHY-2 PCT
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: 202-628-5197
 CC TELEFAX: 202-737-3528
 CC TELEX: 248633
 CC INFORMATION FOR SEQ ID NO: 76:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 298 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS: single
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: peptide
 CC SEQUENCE 298 AA: 33548 MW: 467538 CN:

Query Match 32.4% Score 73: DB 4: Length 298:
 Best Local Similarity 32.3% Pred. No. 1: 97e-01:
 Matches 10: Conservative 10: Mismatches 10: Indels 1: Gaps 1:

Db 50 LFSKAVIALNNGTFLGSPDYVRRVSRIV 79
 QY 1 LFSKLNLIKLNWPLYAGNKKRFKRVSNVI 31

RESULT 4
 ID US-08-537-002A-3 STANDARD: PRT: 963 AA:
 XX XXXXXX
 DE Sequence 3, Application US/08537002A
 CC Sequence 3, Application US/08537002A
 CC Patent No. 5773282
 CC GENERAL INFORMATION:
 CC APPLICANT: TSUSAKI, Keiji
 CC APPLICANT: KUBOTA, Michio
 CC APPLICANT: SUGIMOTO, Toshiyuki

```

CC ZIP: 94111
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/38/484,101B
CC FILING DATE: 07-JUN-1995
CC CLASSIFICATION: 80C
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US94/
CC FILING DATE: 01-JUL-1994
CC CLASSIFICATION: 80C
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/086,555
CC FILING DATE: 31-JUL-1993
CC CLASSIFICATION: 80C
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Trecartin, Richard F.
CC REGISTRATION NUMBER: 31,801
CC REFERENCE/DOCKET NUMBER: A-57515-2/HFT
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 781-1989
CC TELEFAX: (415) 398-3249
CC INFORMATION FOR SEQ ID NO: 48:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 613 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 613 AA; 68332 MW; 1963546 CN;

```

RESULT 6
ID US-C8-484-2019-45
XX
STANDARD: PRT: 6:3 A.M.

| |
|--------|
| A. |
| xx |
| BT |
| xx |
| xxxxxx |

Sequence 46, Application US/08484101B
Patent No. 5824868
GENERAL INFORMATION:
APPLICANT: California Institute of Technology
TITLE OF INVENTION: PLANTS HAVING MODIFIED RESPONSE TO
TITLE OF INVENTION: ETHYLENE
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: Richard F. Trecaire
STREET: 3400 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,101B
FILING DATE: 03 JUN 1995

CC CLASSIFICATION: 800
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: PCT/US94/
CC FILING DATE: 01-JUL-1994
CC CLASSIFICATION: 800
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/086,555
CC FILING DATE: 01-JUL-1993
CC CLASSIFICATION: 800
CC ATTORNEY/AGENT INFORMATION:
CC NAME: FERGUSON, Richard F.
CC REGISTRATION NUMBER: 31,801
CC REFERENCE/DOCKET NUMBER: A-575,5-2/5FT
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 781-1989
CC TELEFAX: (415) 398-3249
CC INFORMATION FOR SEQ ID NO: 46:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 513 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE: 513 AA: 68312 MW: 1963595 CN:
SQ
Query Match: 31.1% Score 70 DB 2: Length 613:
Best Local Similarity 26.7% Pred. No. 3,49e-01:
Matches 8: Conservative 11: Mismatches 11: Indels 0: Gaps 0:
DB 439 LSKNLKLNKMPYAGNKRRIKRV 468
QY 2 FSKNLKLNKMPYAGNKRRIKRVSVI 3:
RESULT 7
ID US-08-749-902-6 STANDARD: PRT: 525 AA
XX
AC xxxxxx
DE Sequence 7, Application US/08749902
XX
Sequence 7, Application US/08749902
CC Patent No. 5385635
CC GENERAL INFORMATION:
CC APPLICANT: Bandman, 1a
CC APPLICANT: Gold, Sui K.
CC APPLICANT: Hillman, Jennifer L.
CC TITLE OF INVENTION: NOVEL HUMAN SERINE/THREONINE
CC TITLE OF INVENTION: PROTEIN KINASES
CC NUMBER OF SEQUENCES: 8
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
CC STREET: 3174 Porter Drive
CC CITY: Palo Alto
CC STATE: CA
CC COUNTRY: US
CC ZIP: 94304
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette
CC COMPUTER: IBM Compatible
CC OPERATING SYSTEM: DOS
CC SOFTWARE: FastSeq Version 2.0
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/749,902
CC FILING DATE: Filed Herewith
CC PRIOR APPLICATION DATA:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Billings, Lucy J.
CC REGISTRATION NUMBER: 36,749
CC REFERENCE/DOCKET NUMBER: PF-0150 US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415-855-0555
CC TELEFAX: 415-845-4166
CC INFORMATION FOR SEQ ID NO: 8:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 525 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:

CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415-855-0555
CC TELEFAX: 415-845-4166
CC INFORMATION FOR SEQ ID NO: 7:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 525 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:
CC LIBRARY: GenBank
CC CL NE: 189508
SQ SEQUENCE 525 AA: 59139 MW: 1455454 CN:
Query Match: 30.7% Score 69: LB 2: Length 525:
Best Local Similarity 34.6% Pred. No. 4,22e-01:
Matches 9: Conservative 10: Mismatches 7: Indels 0: Gaps 0:
DB 302 DKILKCKLNKPYLTQEARLKKL 327
QY 2 FSKNLKLNKMPYAGNKRRIKRV 27
RESULT 8
ID US-08-749-902-6 STANDARD: PRT: 525 AA
XX
AC xxxxxx
DE Sequence 8, Application US/08749902
XX
Sequence 8, Application US/08749902
CC Patent No. 5385635
CC GENERAL INFORMATION:
CC APPLICANT: Bandman, 1a
CC APPLICANT: Gold, Sui K.
CC APPLICANT: Hillman, Jennifer L.
CC TITLE OF INVENTION: NOVEL HUMAN SERINE/THREONINE
CC TITLE OF INVENTION: PROTEIN KINASES
CC NUMBER OF SEQUENCES: 8
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
CC STREET: 3174 Porter Drive
CC CITY: Palo Alto
CC STATE: CA
CC COUNTRY: US
CC ZIP: 94304
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette
CC COMPUTER: IBM Compatible
CC OPERATING SYSTEM: DOS
CC SOFTWARE: FastSeq Version 2.0
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/749,902
CC FILING DATE: Filed Herewith
CC PRIOR APPLICATION DATA:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Billings, Lucy J.
CC REGISTRATION NUMBER: 36,749
CC REFERENCE/DOCKET NUMBER: PF-0150 US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 415-855-0555
CC TELEFAX: 415-845-4166
CC INFORMATION FOR SEQ ID NO: 8:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 525 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:


```

DI XX Sequence 2, Application US/0890980
DE XX Sequence 2, Application US/0890980
CC XX Sequence 2, Application US/0890980
CC XX Patent No. 5998141
CC XX GENERAL INFORMATION:
CC XX APPLICANT: Acton, Susan L.
CC XX TITLE OF INVENTION: SP-B1 NUCLEIC ACIDS AND USES THEREFOR
CC XX NUMBER OF SEQUENCES: 86
CC XX CORRESPONDENCE ADDRESS:
CC XX ADDRESS: FOLEY, ROAG & ELIOT LLP
CC XX STREET: One Post Office Square
CC XX CITY: Boston
CC XX STATE: MA
CC XX COUNTRY: USA
CC XX ZIP: 02109-2170
CC XX COMPUTER READABLE FORM:
CC XX MEDIUM TYPE: Floppy disk
CC XX COMPUTER: IBM PC compatible
CC XX OPERATING SYSTEM: PC-DOS/MS-DOS
CC XX SOFTWARE: Patent's Release #1.0, Version #1.30
CC XX CURRENT APPLICATION DATA: US/08/890,980
CC XX FILING DATE: 10-JUL-1997
CC XX CLASSIFICATION: 435
CC XX ATTORNEY/AGENT INFORMATION:
CC XX NAME: Arnold, Beth E.
CC XX REGISTRATION NUMBER: 35,430
CC XX REFERENCE/DOCKET NUMBER: MIA-005-01
CC XX TELECOMMUNICATION INFORMATION:
CC XX TELEPHONE: 617-832-7000
CC XX TELEFAX: 617-832-7000
CC XX INFORMATION FOR SEQ ID NO: 2:
CC XX SEQUENCE CHARACTERISTICS:
CC XX LENGTH: 509 amino acids
CC XX TYPE: amino acid
CC XX TOPOLOGY: linear
CC XX MOLECULE TYPE: protein
CC XX SEQUENCE 509 AA: 56973 MW: 1456590 CN:

Query Match 30.28; Score 68; DB 2; Length 509;
Best Local Similarity 29.08; Pred. No. 5,09e-01;
Matches 9; Conservative 11; Mismatches 9; Indels 2; Gaps 2;

DE 150 NKEPTIKIMILIAPIIDGKPAFNRIVGEIM 180
QY 13311111111111111111111111111111
3 SKNLIKNNKL-YIAINKKRIKR-VSNVI 31

RESULT 12
ID US-08-752-207B-8 STANDARD; PRT: 605 AA.
XX
AC xxxxxx
DE
DT
DX
XX Sequence 8, Application US/08752307B
DE Sequence 8, Application US/08752307B
CC Sequence 8, Application US/08752307B
CC Patent No. 5952-71
CC XX GENERAL INFORMATION:
CC XX APPLICANT: McCarthy, Sean A.
CC XX APPLICANT: Gearing, David P.
CC XX APPLICANT: Levinson, Douglas A.
CC XX TITLE OF INVENTION: METHOD FOR IDENTIFYING GENES
CC XX TITLE OF INVENTION: ENCODING NOVEL SECRETED OR MEMBRANE-ASSOCIATED PROTEIN
CC XX NUMBER OF SEQUENCES: 14
CC XX CORRESPONDENCE ADDRESS:
CC XX ADDRESS: Fish & Richardson, P.C.
CC XX STREET: 225 Franklin Street
CC XX CITY: Boston
CC XX STATE: MA

```

```

CC COUNTRY: US
CC ZIP: 02110-2804
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette
CC COMPUTER: IBM Compatible
CC OPERATING SYSTEM: Windows95
CC SOFTWARE: FastSeq for Windows Version 2.0
CC CURRENT APPLICATION DATA: US/08/752,07B
CC APPLICATION NUMBER: 08/06/752,07B
CC FILING DATE: 19-NOV-1996
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Meikiejohn, Ph
CC REGISTRATION NUMBER: 283
CC REFERENCE/DOCKET NUMBER: 09404/020001
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 617-542-5670
CC TELEFAX: 617-542-8966
CC TELEX: 200154
CC INFORMATION FOR SEQ ID NO: 8:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 605 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 605 AA: 68102 MW: 1847099 CN:

Query Match 30.28; Score 68; DB 2; Length 605;
Best Local Similarity 33.38; Pred. No. 5,09e-01;
Matches 10; Conservative 11; Mismatches 6; Indels 3; Gaps 2;

DE 164 FPDLSYRWLNEFPVETMDKRRFVSUTN 193
QY 11111111111111111111111111111111
2 FSKNLNIK--LN-MPLSYAGNKKRRFKRVS 28

RESULT 13
ID US-08-040-741-6 STANDARD; PRT: 1018 AA.
XX
AC xxxxxx
DE
DT
DX
XX Sequence 6, Application US/08040741
CC Sequence 6, Application US/08040741
CC Patent No. 6017695
CC XX GENERAL INFORMATION:
CC XX APPLICANT: Reid, Robert A.
CC XX APPLICANT: Hemperly, John J.
CC XX TITLE OF INVENTION: Human Cell Adhesion Molecule and Nucleic
CC XX TITLE OF INVENTION: Acid Sequences
CC XX NUMBER OF SEQUENCES: 6
CC XX CORRESPONDENCE ADDRESS:
CC XX ADDRESSEE: Richard J. Podrick, Becton Dickinson
CC XX ADDRESS: and Company
CC XX STREET: One Becton Drive
CC XX CITY: Franklin Lakes
CC XX STATE: NJ
CC XX COUNTRY: USA
CC XX ZIP: 07417
CC XX COMPUTER READABLE FORM:
CC XX MEDIUM TYPE: Floppy disk
CC XX COMPUTER: IBM PC compatible
CC XX OPERATING SYSTEM: PC-DOS/MS-DOS
CC XX SOFTWARE: Patent In Release #1.0, Version #1.25
CC XX CURRENT APPLICATION DATA:
CC XX APPLICATION NUMBER: US/08/040,741
CC XX FILING DATE: 19930326
CC XX CLASSIFICATION: 436

```


CC TITLE OF INVENTION: Acid Sequences
CC NUMBER OF SEQUENCES: 6
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Richard J. Rodrick, Becton Dickinson and
CC ADDRESSEE: Company
CC STREET: One Becton Drive
CC CITY: Franklin Lakes
CC STATE: NJ
CC COUNTRY: US
CC ZIP: 07417
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent in Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/408,420A
CC FILING DATE:
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Fugit, Donna R.
CC REGISTRATION NUMBER: 32,135
CC REFERENCE/DOCKET NUMBER: P-2630
CC INFORMATION FOR SEQ ID NO: 6:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1018 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC ORIGINAL SOURCE:
CC ORGANISM: Homo sapiens
CC FEATURE:
CC NAME/KEY: Disulfide-bond
CC LOCATION: 45..94
CC FEATURE:
CC NAME/KEY: Disulfide-bond
CC LOCATION: 138..191
CC FEATURE:
CC NAME/KEY: Disulfide-bond
CC LOCATION: 243..290
CC FEATURE:
CC NAME/KEY: Disulfide-bond
CC LOCATION: 332..371
CC FEATURE:
CC NAME/KEY: Disulfide-bond
CC LOCATION: 416..464
CC FEATURE:
CC NAME/KEY: Disulfide-bond
CC LOCATION: 535..563
CC FEATURE:
CC NAME/KEY: Domain
CC LOCATION: 707..760
CC OTHER INFORMATION: /label= FLR
CC OTHER INFORMATION: /note= "conserved core of fibronectin type
CC OTHER INFORMATION: III-like repeat"
CC FEATURE:
CC NAME/KEY: Domain
CC LOCATION: 809..857
CC OTHER INFORMATION: /label= FLR
CC OTHER INFORMATION: /note= "conserved core of fibronectin type
CC OTHER INFORMATION: III-like repeat"
CC FEATURE:
CC NAME/KEY: Domain
CC LOCATION: 905..952
CC OTHER INFORMATION: /label= FLR
CC OTHER INFORMATION: /note= "conserved core of fibronectin type
CC OTHER INFORMATION: III-like repeat"

CC CC
CC NAME/KEY: Modified-site
CC LOCATION: 188
CC OTHER INFORMATION: /label= ASN-glycos
CC OTHER INFORMATION: /note= "potential site of ASN-linked
CC OTHER INFORMATION: glycosylation"
CC FEATURE:
CC NAME/KEY: Modified-site
CC LOCATION: 238
CC OTHER INFORMATION: /label= ASN-glycos
CC OTHER INFORMATION: /note= "potential site of ASN-linked
CC OTHER INFORMATION: glycosylation"
CC FEATURE:
CC NAME/KEY: Modified-site
CC LOCATION: 318
CC OTHER INFORMATION: /label= ASN-glycos
CC OTHER INFORMATION: /note= "potential site of ASN-linked
CC OTHER INFORMATION: glycosylation"
CC FEATURE:
CC NAME/KEY: Modified-site
CC LOCATION: 437
CC OTHER INFORMATION: /label= ASN-glycos
CC OTHER INFORMATION: /note= "potential site of ASN-linked
CC OTHER INFORMATION: glycosylation"
CC FEATURE:
CC NAME/KEY: Modified-site
CC LOCATION: 453
CC OTHER INFORMATION: /label= ASN-glycos
CC OTHER INFORMATION: /note= "potential site of ASN-linked
CC OTHER INFORMATION: glycosylation"
CC FEATURE:
CC NAME/KEY: Modified-site
CC LOCATION: 474
CC OTHER INFORMATION: /label= ASN-glycos
CC OTHER INFORMATION: /note= "potential site of ASN-linked
CC OTHER INFORMATION: glycosylation"
CC FEATURE:
CC NAME/KEY: Modified-site
CC LOCATION: 501
CC OTHER INFORMATION: /label= ASN-glycos
CC OTHER INFORMATION: /note= "potential site of ASN-linked
CC OTHER INFORMATION: glycosylation"
CC FEATURE:
CC NAME/KEY: Modified-site
CC LOCATION: 571
CC OTHER INFORMATION: /label= ASN-glycos
CC OTHER INFORMATION: /note= "potential site of ASN-linked
CC OTHER INFORMATION: glycosylation"
CC FEATURE:
CC NAME/KEY: Modified-site
CC LOCATION: 913
CC OTHER INFORMATION: /label= ASN-glycos
CC OTHER INFORMATION: /note= "potential site of ASN-linked
CC OTHER INFORMATION: glycosylation"
CC FEATURE:
CC NAME/KEY: Modified-site
CC LOCATION: 913
CC OTHER INFORMATION: /label= ASN-glycos
CC OTHER INFORMATION: /note= "potential site of ASN-linked
CC OTHER INFORMATION: glycosylation"
CC SEQUENCE 1018 AA: 113333 MW: 5493875 CN:
Query Match 30.2% Score 68; DB 1; Length 1018;
Best Local Similarity 33.3%; Pred. No. 5,09e+01;
Matches 10; Conservative 1; Mismatches 6; Indels 3; Gaps 2;
Db 164 FPDLSYRWLLNEFPVFTMDKR:PVSQIN 193
1: 111: 11 1111: 11111: 1
QY 2 FSKNLNIK--LN-MPLYIAGNRKRRFKRVS 28

Search completed: Tue Jun 20 13:34:06 2000
Job time : 6 secs.

[W][O][R][D][S]

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit,
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Mon Jun 19 16:22:15 2000; Maspar time 15.47 Seconds
201.982 Million cell updates/sec
Tabular output not generated.

Title: >US-09-142-524A-4
Description: (1-31) from US09142524A.pap
Perfect Score: 225
Sequence: 1 IFSKNLNKLNMPPLYAGNKRRIKRVSNVI 31

Scoring table: PAM 150
Gap 1:

Searched: 721208 seqs, 100765575 residues
Post-Processing: Minimum Match 0%
Listing first 45 summaries

Database: a-pending
1:PT 2:06 3:U60 4:U7 5:U61 7:U82 8:U83 9:U84A
10:U84B 11:U85 12:U86 13:U87 14:U88 15:U89 16:U90 17:U91
18:U92 19:U93 20:U94 21:U95 22:NEWP 23:NEWJ60 24:NEWU8
25:NEWU9
Statistics: Mean 25.139; Variance 88.013; scale 0.297

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

| SUMMARIES | | | | |
|------------|-------|-------------|--------------|--|
| Result No. | Score | Query Match | Length DB ID | Description Pred. No. |
| 1 | 225 | 100.0 | 31 17 | US-09-142- Sequence 4, Applicatio 5.32e-15 |
| 2 | 223 | 54.7 | 346 3 | US-60-109- Sequence 1, Applicatio 1.41e-03 |
| 3 | 223 | 54.7 | 367 10 | US-08-467- Sequence 95, Applicati 1.41e-03 |
| 4 | 223 | 54.7 | 367 10 | US-08-468- Sequence 95, Applicati 1.41e-03 |
| 5 | 223 | 54.7 | 367 7 | US-08-226- Sequence 95, Applicati 1.41e-03 |
| 6 | 223 | 54.7 | 367 8 | US-08-350- Sequence 95, Applicati 1.41e-03 |
| 7 | 223 | 54.7 | 367 10 | US-08-467- Sequence 95, Applicati 1.41e-03 |
| 8 | 223 | 54.7 | 367 10 | US-08-467- Sequence 95, Applicati 1.41e-03 |
| 9 | 223 | 54.7 | 370 10 | US-08-467- Sequence 97, Applicati 1.41e-03 |
| 10 | 223 | 54.7 | 370 10 | US-08-467- Sequence 97, Applicati 1.41e-03 |
| 11 | 223 | 54.7 | 370 7 | US-08-226- Sequence 97, Applicati 1.41e-03 |
| 12 | 223 | 54.7 | 370 10 | US-08-468- Sequence 97, Applicati 1.41e-03 |
| 13 | 223 | 54.7 | 370 10 | US-08-467- Sequence 97, Applicati 1.41e-03 |
| 14 | 223 | 54.7 | 370 8 | US-08-350- Sequence 97, Applicati 1.41e-03 |
| 15 | 118 | 52.4 | 60 7 | US-08-226- Sequence 62, Applicati 4.74e-03 |
| 16 | 118 | 52.4 | 60 10 | US-08-468- Sequence 62, Applicati 4.74e-03 |
| 17 | 118 | 52.4 | 60 10 | US-08-467- Sequence 62, Applicati 4.74e-03 |
| 18 | 118 | 52.4 | 60 8 | US-08-350- Sequence 62, Applicati 4.74e-03 |
| 19 | 118 | 52.4 | 60 10 | US-08-467- Sequence 62, Applicati 4.74e-03 |
| 20 | 118 | 52.4 | 60 4 | US-07-938- Sequence 62, Applicati 4.74e-03 |

| | | | | | |
|----|-----|------|--------|-----------------|----------------------------------|
| 21 | 118 | 52.4 | 60 10 | US-09-142-524-4 | Sequence 52, Applicatio 4.74e-03 |
| 22 | 118 | 52.4 | 134 17 | US-09-142-524-4 | Sequence 3, Applicatio 4.74e-03 |
| 23 | 118 | 52.4 | 353 11 | US-08-526-524-4 | Sequence 15, Applicatio 4.74e-03 |
| 24 | 118 | 52.4 | 374 4 | US-07-938-524-4 | Sequence 2, Applicatio 4.74e-03 |
| 25 | 118 | 52.4 | 374 4 | US-07-730-524-4 | Sequence 2, Applicatio 4.74e-03 |
| 26 | 118 | 52.4 | 374 8 | US-08-350-524-4 | Sequence 2, Applicatio 4.74e-03 |
| 27 | 118 | 52.4 | 374 10 | US-08-467-524-4 | Sequence 2, Applicatio 4.74e-03 |
| 28 | 118 | 52.4 | 374 10 | US-08-467-524-4 | Sequence 2, Applicatio 4.74e-03 |
| 29 | 118 | 52.4 | 374 4 | US-07-723-524-4 | Sequence 2, Applicatio 4.74e-03 |
| 30 | 118 | 52.4 | 374 4 | US-08-467-524-4 | Sequence 2, Applicatio 4.74e-03 |
| 31 | 118 | 52.4 | 374 7 | US-08-226-524-4 | Sequence 2, Applicatio 4.74e-03 |
| 32 | 118 | 52.4 | 374 10 | US-08-468-524-4 | Sequence 2, Applicatio 4.74e-03 |
| 33 | 113 | 50.2 | 20 7 | US-08-226-524-4 | Sequence 33, Applicati 1.58e-02 |
| 34 | 113 | 50.2 | 20 10 | US-08-467-524-4 | Sequence 33, Applicati 1.58e-02 |
| 35 | 113 | 50.2 | 20 10 | US-08-467-524-4 | Sequence 33, Applicati 1.58e-02 |
| 36 | 113 | 50.2 | 20 8 | US-08-350-524-4 | Sequence 33, Applicati 1.58e-02 |
| 37 | 113 | 50.2 | 20 10 | US-08-468-524-4 | Sequence 33, Applicati 1.58e-02 |
| 38 | 113 | 50.2 | 20 10 | US-08-467-524-4 | Sequence 33, Applicati 1.58e-02 |
| 39 | 87 | 38.7 | 105 17 | US-09-142-524-4 | Sequence 5, Applicatio 6.51e-01 |
| 40 | 82 | 36.4 | 31 17 | US-09-142-524-4 | Sequence 5, Applicatio 1.94e-01 |
| 41 | 80 | 35.6 | 80 17 | US-09-142-524-4 | Sequence 1, Applicatio 3.04e-01 |
| 42 | 80 | 35.6 | 377 3 | US-60-096-524-4 | Sequence 15285, Applic 3.04e-01 |
| 43 | 80 | 35.6 | 377 18 | US-09-248-524-4 | Sequence 15285, Applic 3.04e-01 |
| 44 | 78 | 34.7 | 211 18 | US-09-270-524-4 | Sequence 42685, Applic 4.18e-01 |

ALIGNMENTS

RESULT: 1
ID US-09-142-524-4 STANDARD: PRI: 31 AA.
AC xxxxxx
DI
DT
XX
DE Sequence 4, Application US/09142524
XX
CC Sequence 4, Application US/09142524
CC GENERAL INFORMATION:
CC APPLICANT: Sone, Toshio
CC APPLICANT: Kure, Akimori
CC APPLICANT: Kairiki, Kazuo
CC APPLICANT: Iwama, Akiko
CC APPLICANT: King, Kohsuke
CC TITLE OF INVENTION: Pept. based immunotherapeutic Agent For Treating
CC TITLE OF INVENTION: Allergic Diseases
CC FILE REFERENCE: Sone, Toshio
CC CURRENT APPLICATION NUMBER: US/99/142,524
CC CURRENT FILING DATE: 1999-01-04
CC EARLIER APPLICATION NUMBER: B/90/702
CC EARLIER FILING DATE: 1996-03-10
CC EARLIER APPLICATION NUMBER: PCT/JF97/00740
CC EARLIER FILING DATE: 1997-03-10
CC NUMBER OF SEQ ID NOS: 5
CC SOFTWARE: Patentin ver. 2.0
CC SEQ ID NO 4
CC LENGTH: 31
CC TYPE: PRI
CC ORGANISM: Unknown
CC FEATURE:
CC OTHER INFORMATION: Description of Unknown Organism:peptide
SQ SEQUENCE 31 AA: 3661 MW: 4960 CN:

Query Match 100.0%; core 225; DB 17; Length 31;
Best Local Similarity 100.0%; red. No. 5.32e-15;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 IFSKNLNKLNMPPLYAGNKRRIKRVSNVI 31
Qy 1 IFSKNLNKLNMPPLYAGNKRRIKRVSNVI 31

CC REFERENCE/DOCKET NUMBER: 025.6 US (IMI-028CP2)
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 465-6000
CC TELEFAX: (617) 466-6040
CC INFORMATION FOR SEQ ID NO: 95:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 367 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 367 AA: 39846 MW: 695078 CN:

Query Match 54.7% Score 123. DB 8: Length 367;
Best Local Similarity 75.0% Pred. No. 1.41e-03;
Matches 15: Conservative 4: Mismatches 1: Indels 0: Gaps 0:

DB 92 IFSONNLIKMPLYVAGHK 1-1
QY 1 IFSKNLNIKLMPLYAGNK 20

RESULT 7
ID US-08-467-006-95 STANDARD: PRT: 367 AA:

XX xxxxxx

Sequence 95, Application US/08467006

GENERAL INFORMATION:

APPLICANT: Griffith, Irwin J.

APPLICANT: Pollock, Joanne;

APPLICANT: Bond, Julian F.

APPLICANT: Garman, Richard D.

APPLICANT: Kuo, Mel-Chang;

APPLICANT: Yeung, Siu-mei H.

APPLICANT: Brauer, Andrew;

APPLICANT: Exley, Mark A.

APPLICANT: Powers, Steven P.

TITLE OF INVENTION: Allergic Proteins And Fluids From

TITLE OF INVENTION: Japanese Cedar Pollen

NUMBER OF SEQUENCES: 261

CORRESPONDENCE ADDRESS:

ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.

STREET: 610 Lincoln St

CITY: Waltham

STATE: MA

COUNTRY: USA

ZIP: 02154

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/467,006

FILING DATE: June 6, 1995

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/350,225

FILING DATE: December 6, 1994

ATTORNEY/AGENT INFORMATION:

NAME: Jane E. Remillard

REGISTRATION NUMBER: 38,872

REFERENCE/DOCKET NUMBER: 025.6 USD6 (IMI-028CPD6);

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400

TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 95:

SEQUENCE CHARACTERISTICS:

LENGTH: 367 amino acids

CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 367 AA: 39846 MW: 695078 CN:

Query Match 54.7% Score 123. DB 10: Length 367;
Best Local Similarity 75.0% Pred. No. 1.41e-03;
Matches 15: Conservative 4: Mismatches 1: Indels 0: Gaps 0:

DB 92 IFSONNLIKMPLYVAGHK 1-1
QY 1 IFSKNLNIKLMPLYAGNK 20

RESULT 8
ID US-08-467-697-95 STANDARD: PRT: 367 AA:

XX xxxxxx

Sequence 95, Application US/08467697

GENERAL INFORMATION:

APPLICANT: Griffith, Irwin J.

APPLICANT: Pollock, Joanne;

APPLICANT: Bond, Julian F.

APPLICANT: Garman, Richard D.

APPLICANT: Kuo, Mel-Chang;

APPLICANT: Yeung, Siu-mei H.

APPLICANT: Brauer, Andrew;

APPLICANT: Exley, Mark A.

APPLICANT: Powers, Steven P.

TITLE OF INVENTION: Allergic Proteins And Peptides From

TITLE OF INVENTION: Japanese Cedar Pollen

NUMBER OF SEQUENCES: 261

CORRESPONDENCE ADDRESS:

ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.

STREET: 610 Lincoln St

CITY: Waltham

STATE: MA

COUNTRY: USA

ZIP: 02154

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/467,697

FILING DATE: June 6, 1995

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/350,225

FILING DATE: December 6, 1994

ATTORNEY/AGENT INFORMATION:

NAME: Jane E. Remillard

REGISTRATION NUMBER: 38,872

REFERENCE/DOCKET NUMBER: 025.6 USD4 (IMI-028CPD4)

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400

TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 95:

SEQUENCE CHARACTERISTICS:

LENGTH: 367 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE 367 AA: 39846 MW: 695078 CN:

Query Match 54.7% Score 123. DB 10: Length 367;
Best Local Similarity 75.0% Pred. No. 1.41e-03;


```

RESULT 10
ID US-08-467-697-97 STANDARD: PRT: 370 AA.
XX
XX
AC XXXXXX
CC
CC XX
CC XX
CC DT
CC XX
DE DE
XX XX
XX CC
XX CC
Sequence 97, Application US/08467697
Sequence 97, Application US/08467697
GENERAL INFORMATION:
APPLICANT: Griffith Irwin J.;
APPLICANT: Pollock, Anne;
APPLICANT: Bond, Julian F.;
APPLICANT: Garman, Richard D.;
APPLICANT: Kuo, Mei-Chang;
APPLICANT: Yeung, Siu-mei H.;
APPLICANT: Brauer, Andrew;
APPLICANT: Exley, Mark A.;
APPLICANT: Powers, Steven P.
TITLE OF INVENTION: Allergenic Proteins And Peptides From
TITLE OF INVENTION: Japanese Cedar Pollen
NUMBER OF SEQUENCES: 261
CORRESPONDENCE ADDRESS:
ADDRESS: Immunologic Pharmaceutical Corporation, Inc.
STREET: 610 Lincoln St
CITY: Waltham
STATE: MA
COUNTRY: USA
ZIP: 02154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC Compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467-697
FILING DATE: June 6, 1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/350,225
FILING DATE: December 6, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Jane E. Remillard
REGISTRATION NUMBER: 28,872
REFERENCE/DOCKET NUMBER: C25 & USD4 (IM1-028CPD)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO.: 97:
SEQUENCE CHARACTERISTICS:
LENGTH: 370 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE 370 AA: 4019 MW: 725506 CN;
SQ
Query Match 54.7% Score 123; DB 10; Length 370;
Best Local Similarity 75.0%; Pred.No.1.41e-03;
Matches 15; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
Db 92 IFSONNNIKLKMPLYVAGHK 111
QY 1 IFSKNLNKLNMPLYIAGNK 20
RESULT 11
ID US-08-226-248A-97 STANDARD: PRT: 370 AA.
XX
XX XXXXXX
XX AC
XX XX
XX DT
XX XX

```

```

DE Sequence 97, Application US/08226248A
XX Sequence 97, Application US/08226248A
CC GENERAL INFORMATION:
CC APPLICANT: Griffith, Irwin J.:
CC APPLICANT: Pollock, Joanne:
CC APPLICANT: Bond, Julian F.:
CC APPLICANT: Garman, Richard D:
CC APPLICANT: Kuo, Mei-Chang:
CC APPLICANT: Yeung, Siu-mei H.:
CC APPLICANT: Brauer, Andrew:
CC APPLICANT: Exley, Mark A.:
CC APPLICANT: Powers, Steven P.:
CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
CC TITLE OF INVENTION: Japanese Cedar Pollen
CC NUMBER OF SEQUENCES: 201
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
CC STREET: 610 Lincoln St
CC CITY: Waltham
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02154
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent'n Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/226.248A
CC FILING DATE: April 8, 1994
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/350,225
CC FILING DATE: December 6, 1994
CC APPLICATION NUMBER: 08/226,248
CC FILING DATE: April 8, 1994
CC APPLICATION NUMBER: 07/938,990
CC FILING DATE: September 1, 1992
CC APPLICATION NUMBER: PC/US93/00139
CC FILING DATE: January 15, 1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Darlene A. Vanstone
CC REGISTRATION NUMBER: 35,729
CC REFERENCE/DOCKET NUMBER: 025.5 US (IM-028CP)
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 466-6000
CC TELEFAX: (617) 466-6040
CC INFORMATION FOR SEQ ID NO: 97:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 370 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 370 AA: 4019 MW: 725506 CN:

Query Match 54.7% Score 123 DB 7 Length 370:
Best Local Similarity 75.0% Pred. No. 1,41e-03:
Matches 15: Conservative 4: Mismatches 1: Indels 0: Gaps 0:

DE Sequence 97, Application US/0846894C
XX Sequence 97, Application US/0846894C
CC GENERAL INFORMATION:
CC APPLICANT: Griffith, Irwin J.:
CC APPLICANT: Pollock, Joanne:
CC APPLICANT: Bond, Julian F.:
CC APPLICANT: Garman, Richard D:
CC APPLICANT: Kuo, Mei-Chang:
CC APPLICANT: Yeung, Siu-mei H.:
CC APPLICANT: Brauer, Andrew:
CC APPLICANT: Exley, Mark A.:
CC APPLICANT: Powers, Steven P.:
CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
CC TITLE OF INVENTION: Japanese Cedar Pollen
CC NUMBER OF SEQUENCES: 201
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
CC STREET: 610 Lincoln St
CC CITY: Waltham
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02154
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent'n Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/226.248A
CC FILING DATE: April 8, 1994
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/938,990
CC FILING DATE: September 1, 1992
CC APPLICATION NUMBER: PC/US93/00139
CC FILING DATE: January 15, 1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Darlene A. Vanstone
CC REGISTRATION NUMBER: 35,729
CC REFERENCE/DOCKET NUMBER: 025.5 US (IM-028CP)
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 466-6000
CC TELEFAX: (617) 466-6040
CC INFORMATION FOR SEQ ID NO: 97:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 370 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 370 AA: 4019 MW: 725506 CN:

Query Match 54.7% Score 123 DB 7 Length 370:
Best Local Similarity 75.0% Pred. No. 1,41e-03:
Matches 15: Conservative 4: Mismatches 1: Indels 0: Gaps 0:

DE Sequence 97, Application US/08457006
XX Sequence 97, Application US/08457006
CC GENERAL INFORMATION:
CC APPLICANT: Griffith, Irwin J.:
CC APPLICANT: Pollock, Joanne:

```

CC APPLICANT: Bond, Julian F.;
CC APPLICANT: Garman, Richard D.;
CC APPLICANT: Kuo, Mei-Chang;
CC APPLICANT: Yeung, Siu-mei H.;
CC APPLICANT: Brauer, Andrew;
CC APPLICANT: Exley, Mark A.;
CC APPLICANT: Powers, Steven P.
CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
CC TITLE OF INVENTION: Japanese Cedar Pollen
CC NUMBER OF SEQUENCES: 261
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: ImmuLogic Pharmaceutical Corporation, Inc.
CC STREET: 610 Lincoln St
CC CITY: Waltham
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02154
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/467,006
CC FILING DATE: June 6, 1995
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/350,225
CC FILING DATE: December 6, 1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Case E. Remillard
CC REGISTRATION NUMBER: 38,872
CC REFERENCE/DOCKET NUMBER: 025.6 USD6 (IMI-028CPD6)
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 227-7400
CC TELEFAX: (617) 227-5941
CC INFORMATION FOR SEQ ID NO: 97:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 370 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 370 AA: 40191 MW: 725506 CN:
Query Match 54.7% Score 123 DB 10: Length 370:
Best Local Similarity 75.0% Pred. No. 1,41e-03:
Matches 15: Conservative 4: Mismatches 11: Indels 0: Gaps 0:
DC 92 IFSQNNNIKIKYPLVVAQNK 11
CY 1 IFSKNLNKLNKPLYIAQNK 20
RESULT 14
ID US-08-350-225-97 STANDARD: PRT: 370 AA.
XX XXXXXX
AC
XX
DI
DE
XX
XX
Sequence 97, Application US/08350225
CC
CC Sequence 97, Application US/08350225
CC GENERAL INFORMATION:
CC APPLICANT: Griffith, Irwin J.;
CC APPLICANT: Pollock, Joanne;
CC APPLICANT: Bond, Julian F.;
CC APPLICANT: Garman, Richard D.;
CC APPLICANT: Kuo, Mei-Chang;
CC APPLICANT: Yeung, Siu-mei H.;
CC APPLICANT: Brauer, Andrew;
CC APPLICANT: Exley, Mark A.;
CC APPLICANT: Powers, Steven P.

CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
CC TITLE OF INVENTION: Japanese Cedar Pollen
CC NUMBER OF SEQUENCES: 26:
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: ImmuLogic Pharmaceutical Corporation, Inc.
CC STREET: 610 Lincoln S
CC CITY: Waltham
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02154
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/350,225
CC FILING DATE: December 6, 1994
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/226,248
CC FILING DATE: April 6, 1994
CC APPLICATION NUMBER: 07/938,990
CC FILING DATE: September 11, 1992
CC APPLICATION NUMBER: PCT/US93/00139
CC FILING DATE: January 1, 1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Darlene A. Vaustrine
CC REGISTRATION NUMBER: 3,729
CC REFERENCE/DOCKET NUMBER: 025.6 US (IMI-028CP2)
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 466-6300
CC TELEFAX: (617) 466-6300
CC INFORMATION FOR SEQ ID NO: 97:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 370 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 370 AA: 40191 MW: 725506 CN:
Query Match 54.7% Score 123 DB 8: Length 370:
Best Local Similarity 75.0% Pred. No. 1,41e-03:
Matches 15: Conservative 4: Mismatches 11: Indels 0: Gaps 0:
DC 92 IFSQNNNIKIKYPLVVAQNK 11
CY 1 IFSKNLNKLNKPLYIAQNK 20
RESULT 15
ID US-08-226-248A-62 STANDARD: PRT: 60 AA.
XX XXXXXX
AC
XX
DI
DE
XX
XX
Sequence 62, Application US/08226248A
CC
CC Sequence 62, Application US/08226248A
CC GENERAL INFORMATION:
CC APPLICANT: Griffith, Irwin J.;
CC APPLICANT: Pollock, Joanne;
CC APPLICANT: Bond, Julian F.;
CC APPLICANT: Garman, Richard D.;
CC APPLICANT: Kuo, Mei-Chang;
CC APPLICANT: Yeung, Siu-mei H.;
CC APPLICANT: Brauer, Andrew;
CC APPLICANT: Exley, Mark A.;
CC APPLICANT: Powers, Steven P.
CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
CC TITLE OF INVENTION: Japanese Cedar Pollen
CC NUMBER OF SEQUENCES: 201

```

CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
CC STREET: 610 Lincoln St
CC CITY: Waltham
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02154
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC Compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/225,248A
CC FILING DATE: April 8, 1994
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/938,990
CC FILING DATE: September 1, 1992
CC APPLICATION NUMBER: PCT/US93/00139
CC FILING DATE: January 15, 1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Darlene A. VanStone
CC REGISTRATION NUMBER: 35,729
CC REFERENCE/DOCKET NUMBER: 025.5 US (MI-028CP)
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 465-6000
CC TELEFAX: (617) 466-6040
CC INFORMATION FOR SEQ ID NO: 62:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 60 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC FRAGMENT TYPE: internal
CC SEQUENCE 60 AA: 6644 MW: 19464 CN:

Query Match 52.4% Score 18: D9.7: length 60:
Best Local Similarity 60.0% Pred. No. 4.7 e-03:
Matches 18: Conservative 4: Mismatch 7: Indels 1: Gaps 1:

DB 11 IFSONMNIKKMPYIAGYKT-FDGRGAQV 39
QY 11 FSKNLNLIKMPYIAGNKKRR-KRVSNV 30

Search completed: Mon Jun 19 10:22:45 2000
JCL time : 30 secs

```



```

ACCESSIONS      26-Aug-1999
REFERENCE        J02123: PC2065
AUTHORS         Sene, T.; Komiyama, N.; Shimizu, K.; Sakabe, T.; Morikubo,
                K.; Kinc, K.
JOURNAL         Biochem. Biophys. Res. Commun. (1994) 199:610-625
TITLE           Cloning and sequencing of cDNA coding for Cry j I, a major
                allergen of Japanese cedar pollen.
CROSS-REFERENCES MIMD:9419334
ACCESSION       J02123
MOLECULE_TYPE   mRNA
RESIDUES        1-374 #label SON
CROSS-REFERENCES GB:025544; MID:q49363; PID:d1006086; PID:q493632
EXPERIMENTAL_SOURCE pollen
ACCESSION       PC2065
MOLECULE_TYPE   protein
RESIDUES        22-53:58-81:219-232:236-258:269-307:346-372 #label S02
NOTE            the authors described carbohydrate binding site for
                residue 279
CLASSIFICATION  #superfamily pectate lyase LA759
KEYWORDS        glycoprotein; pollen
FEATURE         1-21 #domain signal sequence #status predicted #label SIGN
                22-374 #product major allergen Cry j 1 (clone PC02-2-2) #status
                159,191,203,354 #binding_site carbohydrate (Asn) (covalent) #status
                predicted
SUMMARY         #length 374 #molecular_weight 43645 #checksum 2920
Query Match    52.4%; Score 118; DB 2; Length 374;
Best Local Similarity 60.0%; Pred. NO. 5.55e-07;
Matches 18; Conservative 4; Mismatches 7; Indels 1; Gaps 1;

DB 92 IFSCNNIKLKMPLYIAGYKT-FDGRGVV 120
QY 1 IFSCNNIKLKMPLYIAGNKRPFIRKVSNV 30

RESULT 3
ENTRY #16992 #type complete
TITLE ethylene receptor homolog - apple tree
ORGANISM #normal_name Malus domestica #common_name apple tree
DATE 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change
ACCESSIONS #16992
REFERENCE Z16395
AUTHORS Chen, H.H.; Chertov, Y.Y.; Yang, S.F.; Shaw, J.F.
JOURNAL Plant Physiology. (1998) 117:1125
TITLE Isolation and characterization of a broccoli cDNA (Accession
        No. AF047477) encoding an ERS-type ethylene receptor
        (P0898-123).
ACCESSION #16992
STATUS preliminary; translated from GB/EMBL/DBJ
MOLECULE_TYPE mRNA
RESIDUES 1-741 #label LEE
CROSS-REFERENCES EMBL:AF032448; MID:q341050; PID:q3411051
EXPERIMENTAL_SOURCE cultivar Granny Smith; ripening fruit
GENETICS
NOTE ERI
SUMMARY #length 741 #molecular_weight 82967 #checksum 4504
Query Match 37.3%; Score 84; DB 2; Length 741;
Best Local Similarity 33.3%; Pred. NO. 1.97e-01;
Matches 10; Conservative 10; Mismatches 10; Indels 0; Gaps 0;

DB 439 LSVTLNIAADLPYVIGDERRLMOTILNV 468
QY 2 FSKNLNLIKLNPLYIAGNKRPFIRKVSNI 31

RESULT 4
ENTRY #52826 #type complete
TITLE hypothetical protein YMR066w - yeast (Saccharomyces

```

```

cerevisiae)
ALTERNATE_NAMES hypothetical protein YMR916.05
ORGANISM #formal_name Saccharomyces cerevisiae
DATE 19-May-1995 #sequence_revision 01-Sep-1995 #text_change
21-Nov-1997
ACCESSIONS S52826
REFERENCE S52814
AUTHORS Pearson, D.; Bowman, S.
SUBMISSION submitted to the EMBL Data Library, April 1995
ACCESSION S52826
MOLECULE_TYPE DNA
RESIDUES 1-898 #label PEA
CROSS-REFERENCES EMBL:Z48952; MID:q763008; PID:q763013; MIPS:YRP_PEA
EXPERIMENTAL_SOURCE strain AB972
GENETICS
SUMMARY #map_position 13R
        #length 898 #molecular_weight 104747 #checksum 7641
Query Match 36.9%; Score 83; DB 2; Length 898;
Best Local Similarity 37.0%; Pred. NO. 2.78e-01;
Matches 10; Conservative 11; Mismatches 5; Indels 1; Gaps 1;

DB 785 DLNIDYDLSMYLKGKRCYLROISNT 811
QY 5 NLNKLNMPLYIA-GNKRPFIRKVSNV 30

RESULT 5
ENTRY #53852 #type complete
TITLE ribosomal protein L5 - Acanthamoeba castellanii (Pilot edition
        (SGC5))
ORGANISM #formal_name mitochondrion Acanthamoeba castellanii
DATE 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change
26-Jun-1999
ACCESSIONS S53852
REFERENCE S53825
AUTHORS Burger, G.; Pilsch, J.; Loneragan, K.M.; Gray, M.W.
JOURNAL J. Mol. Biol. (1995) 245:522-537
TITLE The mitochondrial DNA of the amoeboid protozoan, Acanthamoeba
        castellanii: complete sequence, gene content and amino acid
        organization.
CROSS-REFERENCES MIMD:95147275
ACCESSION S53852
STATUS nucleic acid sequence not shown; translation not shown
MOLECULE_TYPE DNA
RESIDUES 1-177 #label RRP
CROSS-REFERENCES GB:013494; MID:q66247; PID:q66247
EXPERIMENTAL_SOURCE strain Acanthamoeba castellanii
NOTE the nucleotide sequence was submitted to the EMBL Data
        Library, July 1994
GENETICS
NOTE mitochondrion
GENETIC_CODE SGC6
KEYWORDS mitochondrion
SUMMARY #length 177 #molecular_weight 21764 #checksum 4772
Query Match 36.0%; Score 81; DB 2; Length 177;
Best Local Similarity 50.0%; Pred. NO. 5.49e-01;
Matches 13; Conservative 6; Mismatches 4; Indels 3; Gaps 3;

DB 46 RSLKPFIKLYLWLYIIS-NQKPFIRKVS 70
QY 4 KNLN-IKLN-WPLYIAGNKRPFIRKVS 27

RESULT 6
ENTRY #52611 #type complete
TITLE pectate lyase (EC 4.2.2.2) - common tobacco
ORGANISM #formal_name Nicotiana tabacum #common_name common tobacco
DATE 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change
29-Sep-1999
ACCESSIONS S26211; S26212; S21933; S22753; S22754
REFERENCE S26211

```


##cross-references GB:AE000706; NID:q2983327; PID:q2983337; GB:AE000657
 ##experimental_source strain VF5

GENETICS

##gene

##summary

##length 480 #molecular-weight 55991 #checksum 8726

Query Match 34.2% Score 77; DB 2; Length 480;

Best Local Similarity 32.1% Pred. No. 2.08e+00;

Matches 9; Conservative 10; Mismatches 9; Indels 0; Gaps 0;

Db 377 IFSTAYQISAEIAFYVQNPRTYTHVN 404

||||| 1 ||||| 10 ||| 11

QY 1 IFSKNLNKLNNPLVIAGNKRERFKRVS 28

RESULT 12

ENTRY

##type complete

##title

##organism

##date

##accessions

##reference

##authors

##submission

##description

##accession

##status

##molecule_type DNA

##residues 1-532 #label FAV

##cross-references EMBL:J23515; NID:q746492; PID:q746497;

##experimental_source strain Bristol N2

GENETICS

##gene

##introns

##summary

##length 532 #molecular-weight 59809 #checksum 4155

Query Match

Best Local Similarity 35.5% Score 76; DB 2; Length 532;

Matches 11; Conservative 11; Mismatch 5; Indels 2; Gaps 2;

Db 180 JFSANK-KLVSPVLC-NRR-LFIRHVTDIV 208

||||| 1 ||||| 1 ||||| 11 ||||| 11

QY 1 IFSKNLNKLNNPLVIAGNKRERFKRVS 31

RESULT 13

ENTRY

##type complete

##title

##organism

##date

##accessions

##reference

##authors

Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann,

R.D.; Sutton, G.G.; Blake, J.A.; FitzGerald, L.M.; Clayton,

R.A.; Gocayne, J.D.; Kerlavage, A.R.; Dougherty, B.A.;

Tomb, J.F.; Adams, M.D.; Reich, C.L.; Overbeek, R.;

Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;

Scott, J.L.; Geoghegan, N.S.M.; Weidman, J.F.; Fuhrmann,

J.L.; Nguyen, D.; Utterback, T.R.; Kelley, J.M.; Peterson,

J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts,

K.M.; Fraser, C.M.; Kaine, B.P.; Borodovsky, M.; Klenk,

H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.R.; Venter, J.C.

Science (1996) 273:1058-1073

Complete genome sequence of the methanogenic archaeon,

Methanococcus jannaschii.

##cross-references MUID:96337999

##accession

##status

##molecule_type DNA

preliminary: nucleic acid sequence not shown;

translation not shown;

##residues 1-569 #label BUL
 ##cross-references GB:U67545; GB:L77117; NID:q1591480; PID:q1499364;
 TIGR:M1022

GENETICS

##map_position REV953676-951967

##summary

##length 569 #molecular-weight 6558; #checksum 4248

Query Match 33.8% Score 75; DB 2; Length 569;

Best Local Similarity 32.7% Pred. No. 2.89e+00;

Matches 11; Conservative 8; Mismatches 6; Indels 2; Gaps 2;

Db 102 FKHVDLIEEDADYVLIADGRRLLIER 128

||||| 1 ||||| 11 ||| 1

QY 2 FSKNLN-IKL-NMPLYIAGNKRERFKR 26

RESULT 14

ENTRY

##type complete

##title

##organism

##date

##accessions

##reference

##authors

##submission

##description

##accession

##molecule_type genomic RNA

##residues 1-3140 #label MA1

##cross-references EMBL:X81083; NID:q531731; PID:CAA56974 1;

##experimental_source isolate PVV-SC

##classification

##keywords

##feature

1-3046

309-326

767-1176

117-1168

169-1803

1253-1526

1253-1260

1338-1343

1342-1345

1804-1856

1857-2292

1857-2049

2293-2810

2811-3140

1919

##product protein F1 #status predicted #label P1P1

##product protein F2 #status predicted #label P1P1

##product protein F3 #status predicted #label P1P1

##product protein F4 #status predicted #label P1P1

##product protein F5 #status predicted #label P1P1

##product protein F6 #status predicted #label P1P1

##product protein F7 #status predicted #label P1P1

##product protein F8 #status predicted #label P1P1

##product protein F9 #status predicted #label P1P1

##product protein F10 #status predicted #label P1P1

##product protein F11 #status predicted #label P1P1

##product protein F12 #status predicted #label P1P1

##product protein F13 #status predicted #label P1P1

##product protein F14 #status predicted #label P1P1

##product protein F15 #status predicted #label P1P1

##product protein F16 #status predicted #label P1P1

##product protein F17 #status predicted #label P1P1

##product protein F18 #status predicted #label P1P1

##product protein F19 #status predicted #label P1P1

##product protein F20 #status predicted #label P1P1

##product protein F21 #status predicted #label P1P1

##product protein F22 #status predicted #label P1P1

##product protein F23 #status predicted #label P1P1

##product protein F24 #status predicted #label P1P1

##product protein F25 #status predicted #label P1P1

##product protein F26 #status predicted #label P1P1

##product protein F27 #status predicted #label P1P1

##product protein F28 #status predicted #label P1P1

##product protein F29 #status predicted #label P1P1

##product protein F30 #status predicted #label P1P1

Search completed: Mon Jan 19 16:00:43 2000
Job time : 10 secs.

[W][O][R][D][S] [T][M]

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Mon Jun 19 15:59:18 2000; MasPar time 3.43 seconds
Tabular output not generated. 274,849 Million cell updates/sec

Title: >US-09-142-524A-4
Description: (1-31) from US09142524A.pep
Perfect Score: 225
Sequence: 1:FSKNVNLKMLNPLYIAGNRRRIKRVSNV: 31

Scoring table: PAM 150
Gap 11

Searched: 83957 seqs, 30454973 residues
Post-processing: Minimum Match: 0%
Listing first 45 summaries

Database: swiss-pro:38
::swissprot

Statistics: Mean 34.404; Variance 55.946; scale 0.615

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | BP | ID | Description | Prod. No. |
|------------|-------|---------------|--------|----|------------|------------------------|-----------|
| 1 | 118 | 52.4 | 374 | 1 | SBP_CRYJA | SUGI BASIC PROTEIN PRE | 2 90e-08 |
| 2 | 83 | 36.9 | 598 | 1 | YMW5_YEAST | HYPOTHETICAL 104.7 KD | 6 42e-02 |
| 3 | 81 | 36.0 | 127 | 1 | RM05_ACACA | MITOCHONDRIAL 60S RIBO | 1 34e-01 |
| 4 | 81 | 36.0 | 397 | 1 | PEL1_TORAC | PECTATE LYASE PRECURSO | 1 34e-01 |
| 5 | 81 | 36.0 | 3140 | 1 | POLG_PPVRA | GENOME POLYPROTEIN [CO | 1 34e-01 |
| 6 | 81 | 36.0 | 3141 | 1 | POLG_PPVD | GENOME POLYPROTEIN [CO | 1 34e-01 |
| 7 | 78 | 34.7 | 143 | 1 | YEH7_YEAST | VERY HYPOTHETICAL 16.3 | 4 06e-01 |
| 8 | 76 | 33.8 | 532 | 1 | YRS5_GAEEL | HYPOTHETICAL 59.8 KD P | 8 38e-01 |
| 9 | 76 | 33.8 | 569 | 1 | YAZ2_METJA | HYPOTHETICAL PROTEIN M | 8 38e-01 |
| 10 | 75 | 33.3 | 347 | 1 | RH31_YEAST | DNA DAMAGE TOLERANCE P | 1 20e+00 |
| 11 | 74 | 32.9 | 90 | 1 | AX68_SOTBN | AUXIN-INDUCED PROTEIN | 1 71e+00 |
| 12 | 74 | 32.9 | 92 | 1 | ARG7_PHAAR | INDOLE-3-ACETIC ACID | 1 71e+00 |
| 13 | 74 | 32.9 | 189 | 1 | Y573_METJA | HYPOTHETICAL PROTEIN M | 1 71e+00 |
| 14 | 73 | 32.4 | 343 | 1 | RTA_RAT | PROBABLE G PROTEIN-COU | 2 42e+00 |
| 15 | 73 | 32.4 | 3125 | 1 | POLG_PPVNA | GENOME POLYPROTEIN [CO | 2 42e+00 |
| 16 | 72 | 32.0 | 238 | 1 | RM08_YEAST | MITOCHONDRIAL 60S RIBO | 3 43e+00 |
| 17 | 72 | 32.0 | 309 | 1 | VIRB_SHIEL | VIROLENCE REGUCON TRAN | 3 43e+00 |
| 18 | 72 | 32.0 | 3140 | 1 | POLG_PPVSX | GENOME POLYPROTEIN [CO | 3 43e+00 |
| 19 | 71 | 31.6 | 136 | 1 | R17A_YEAST | 40S RIBOSOMAL PROTEIN | 4 83e+00 |
| 20 | 71 | 31.6 | 202 | 1 | YC63_HUMAN | HYPOTHETICAL PROTEIN K | 4 83e+00 |
| 21 | 71 | 31.6 | 255 | 1 | MODE_HAEIN | MOLYBDENUM TRANSPORT P | 4 83e+00 |
| 22 | 71 | 31.6 | 506 | 1 | R1RL_HCMVA | RIBONUCLEOSIDE-DIPHOSP | 4 83e+00 |
| 23 | 70 | 31.1 | 327 | 1 | Y013_NPVAC | HYPOTHETICAL 38.7 KDP | 6 78e+00 |

| ALIGNMENTS | | | | | | | | | |
|------------|----|-----------|--------|----|------------|------------------------|----------|--------------|----|
| RESULT | | 1 | | ID | | STANDARD: | | PRG: 374 AA. | |
| AC | AD | SBP_CRYJA | PI8632 | AC | AD | SBP_CRYJA | PI8632 | AC | AD |
| 24 | 70 | 31.1 | 746 | 1 | TAGF_BACSU | TEICHOIC ACID BIOSYNTH | 6.78e+00 | | |
| 25 | 70 | 31.1 | 1518 | 1 | POLG_PPVEA | GENOME POLYPROTEIN [CC | 6.78e+00 | | |
| 26 | 69 | 30.7 | 247 | 1 | YACE_ECOLI | HYPOTHETICAL 28.3 KD P | 9.49e+00 | | |
| 27 | 69 | 30.7 | 333 | 1 | IF2B_RABIT | EUKARYOTIC TRANSLATION | 9.49e+00 | | |
| 28 | 69 | 30.7 | 333 | 1 | IF2B_HUMAN | EUKARYOTIC TRANSLATION | 9.49e+00 | | |
| 29 | 69 | 30.7 | 351 | 1 | IGNA_PSES2 | INCREASED GLYPHOSATE R | 9.49e+00 | | |
| 30 | 69 | 30.7 | 502 | 1 | KSO_RAT | RIBOSOMAL PROTEIN S6 K | 9.49e+00 | | |
| 31 | 69 | 30.7 | 502 | 1 | KSO_HUMAN | RIBOSOMAL PROTEIN S6 K | 9.49e+00 | | |
| 32 | 69 | 30.7 | 686 | 1 | MX1A_SHIEL | MXIA PROTEIN (VIRH PRO | 9.49e+00 | | |
| 33 | 68 | 30.2 | 376 | 1 | Y501_MYCTU | HYPOTHETICAL 41.1 KD P | 1.32e+01 | | |
| 34 | 68 | 30.2 | 395 | 1 | Y164_METJA | HYPOTHETICAL PROTEIN M | 1.32e+01 | | |
| 35 | 68 | 30.2 | 899 | 1 | YM23_YEAST | HYPOTHETICAL 103 C KD | 1.32e+01 | | |
| 36 | 68 | 30.2 | 1018 | 1 | CONT_HUMAN | CONTACTIN PRECURSOR (G | 1.32e+01 | | |
| 37 | 68 | 30.2 | 1020 | 1 | CONT_MOUSE | CONTACTIN PRECURSOR (M | 1.32e+01 | | |
| 38 | 67 | 29.8 | 92 | 1 | AX13_SOYBN | AUXIN-INDUCED PROTEIN | 1.64e+01 | | |
| 39 | 67 | 29.8 | 181 | 1 | RL11_MEDSA | 60S RIBOSOMAL PROTEIN | 1.84e+01 | | |
| 40 | 67 | 29.8 | 184 | 1 | RL12_ARATH | 60S RIBOSOMAL PROTEIN | 1.84e+01 | | |
| 41 | 67 | 29.8 | 256 | 1 | BLAB_BACCE | BETA-LACTAMASE, TYPE I | 1.84e+01 | | |
| 42 | 67 | 29.8 | 392 | 1 | MP14_AMBAR | POLLEN ALLERGEN AMB A | 1.84e+01 | | |
| 43 | 67 | 29.8 | 427 | 1 | Y428_METJA | HYPOTHETICAL PROTEIN M | 1.84e+01 | | |
| 44 | 67 | 29.8 | 1039 | 1 | YLA6_GAEEL | HYPOTHETICAL 121.8 KD | 1.84e+01 | | |
| 45 | 67 | 29.8 | 1092 | 1 | DPOG_DROME | DNA POLYMERASE DELTA C | 1.84e+01 | | |
| ALIGNMENTS | | | | | | | | | |
| RESULT | | 1 | | ID | | STANDARD: | | PRG: 374 AA. | |
| AC | AD | SBP_CRYJA | PI8632 | AC | AD | SBP_CRYJA | PI8632 | AC | AD |
| 24 | 70 | 31.1 | 746 | 1 | TAGF_BACSU | TEICHOIC ACID BIOSYNTH | 6.78e+00 | | |
| 25 | 70 | 31.1 | 1518 | 1 | POLG_PPVEA | GENOME POLYPROTEIN [CC | 6.78e+00 | | |
| 26 | 69 | 30.7 | 247 | 1 | YACE_ECOLI | HYPOTHETICAL 28.3 KD P | 9.49e+00 | | |
| 27 | 69 | 30.7 | 333 | 1 | IF2B_RABIT | EUKARYOTIC TRANSLATION | 9.49e+00 | | |
| 28 | 69 | 30.7 | 333 | 1 | IF2B_HUMAN | EUKARYOTIC TRANSLATION | 9.49e+00 | | |
| 29 | 69 | 30.7 | 351 | 1 | IGNA_PSES2 | INCREASED GLYPHOSATE R | 9.49e+00 | | |
| 30 | 69 | 30.7 | 502 | 1 | KSO_RAT | RIBOSOMAL PROTEIN S6 K | 9.49e+00 | | |
| 31 | 69 | 30.7 | 502 | 1 | KSO_HUMAN | RIBOSOMAL PROTEIN S6 K | 9.49e+00 | | |
| 32 | 69 | 30.7 | 686 | 1 | MX1A_SHIEL | MXIA PROTEIN (VIRH PRO | 9.49e+00 | | |
| 33 | 68 | 30.2 | 376 | 1 | Y501_MYCTU | HYPOTHETICAL 41.1 KD P | 1.32e+01 | | |
| 34 | 68 | 30.2 | 395 | 1 | Y164_METJA | HYPOTHETICAL PROTEIN M | 1.32e+01 | | |
| 35 | 68 | 30.2 | 899 | 1 | YM23_YEAST | HYPOTHETICAL 103 C KD | 1.32e+01 | | |
| 36 | 68 | 30.2 | 1018 | 1 | CONT_HUMAN | CONTACTIN PRECURSOR (G | 1.32e+01 | | |
| 37 | 68 | 30.2 | 1020 | 1 | CONT_MOUSE | CONTACTIN PRECURSOR (M | 1.32e+01 | | |
| 38 | 67 | 29.8 | 92 | 1 | AX13_SOYBN | AUXIN-INDUCED PROTEIN | 1.64e+01 | | |
| 39 | 67 | 29.8 | 181 | 1 | RL11_MEDSA | 60S RIBOSOMAL PROTEIN | 1.84e+01 | | |
| 40 | 67 | 29.8 | 184 | 1 | RL12_ARATH | 60S RIBOSOMAL PROTEIN | 1.84e+01 | | |
| 41 | 67 | 29.8 | 256 | 1 | BLAB_BACCE | BETA-LACTAMASE, TYPE I | 1.84e+01 | | |
| 42 | 67 | 29.8 | 392 | 1 | MP14_AMBAR | POLLEN ALLERGEN AMB A | 1.84e+01 | | |
| 43 | 67 | 29.8 | 427 | 1 | Y428_METJA | HYPOTHETICAL PROTEIN M | 1.84e+01 | | |
| 44 | 67 | 29.8 | 1039 | 1 | YLA6_GAEEL | HYPOTHETICAL 121.8 KD | 1.84e+01 | | |
| 45 | 67 | 29.8 | 1092 | 1 | DPOG_DROME | DNA POLYMERASE DELTA C | 1.84e+01 | | |

RESULT 1
ID SBP_CRYJA STANDARD: PRG: 374 AA.
AC P18632:
DI 01-NOV-1990 (Rel. 16, Created)
DI 01-NOV-1995 (Rel. 32, Last sequence update)
DI 15-DEC-1998 (Rel. 37, Last annotation update)
DE SUGI BASIC PROTEIN PRECURSOR (SBP) (MAJOR ALLERGEN CRY J 1) (CRY J 1)
OS Cryptomeria japonica (Japanese cedar)
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta
OC euphyllophytes: Spermatophyta: Coniferopsida: Coniferales:
OC Taxodiaceae: Cryptomeria
RN 11]
SEQUENCE FROM N.A. AND PAK. AL. SEQUENCE.
RC TISSUE-POLLEN:
RX MEDLINE: 94183234
RA Sone T., Komiyama N., Shimizu K., Kusakabe T., Morikubo K.,
Kino K.:
RT "Cloning and sequencing of cDNA coding for Cry J 1, a major allergen
of Japanese cedar pollen."
RL Biochem. Biophys. Res. Commun. 199 619-625(1994).
RN 12]
SEQUENCE FROM N.A.
RC TISSUE-POLLEN:
RA Namba M., Karose M., Torigoe K., Fukuda S., Kurimoto M.,
Submitted (201-1994) to the EMBL/GenBank/DBS databases.
RN 13]
SEQUENCE OF 22-41.
RC TISSUE-POLLEN:
RX MEDLINE: 89031257.
RA Tanai M., Ando S., Usui M., Kurimoto M., Sakaguchi M., Inouye S.,
Matuhasi T.:
RT "N-terminal amino acid sequence of a major allergen of Japanese cedar
pollen (Cry j 1)."
RL FEBS Lett. 239:329-332(1988).
RN 14]
CARBOHYDRATES.
RX MEDLINE: 95003748.
RA Hilkata A., Matsumoto T., Kojima K., Ogawa H.:
RT "Antigenicity of the oligosaccharide moiety of the Japanese cedar
(Cryptomeria japonica) pollen allergen, Cry j1."
RL Int. Arch. Allergy Immunol. 105:198-202(1994).
CC 1- PTM: CONTAINS FUCOSE/XULOSE-CONTAINING N-LINKED OLIGOSACCHARIDES.
CC 1- DISEASE: THIS PROTEIN IS THE MAJOR ALLERGEN OF JAPANESE CEDAR
POLLEN. THE MOST COMMON POLLEN ALLERGEN IN JAPAN.
CC 1- MISCELLANEOUS: THE SEQUENCE OF CRY J 1 FORM A IS SHOWN HERE. FORM
B DIFFERS IN SIX POSITIONS.
CC 1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.

```

CC      AMB A 1/AMB A 11/CRY 3 1 SUBFAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: D26544; BAA05542.1;
CC      EMBL: D26545; BAA05543.1;
CC      EMBL: D34639; BAA07020.1;
CC      PIR: A44773; A44773;
CC      PIR: PFC0344; pcclyase;
CC      PRINTS: PRO087; AMBALLERGEN;
CC      KW: Allergen; Glycoprotein; Multigene family; Signal;
CC      FT: SIGNAL 1 21
CC      FT: CHAIN 22 374
CC      FT: SIGNAL 22 374
CC      FT: VARIANT 12 12
CC      FT: VARIANT 143 143
CC      FT: VARIANT 202 202
CC      FT: VARIANT 221 221
CC      FT: VARIANT 221 221
CC      FT: VARIANT 358 358
CC      FT: VARIANT 361 361
CC      FT: VARIANT 158 158
CC      FT: CARBOHYD 191
CC      FT: CARBOHYD 293
CC      FT: POTENTIAL
CC      FT: POTENTIAL
CC      SEQUENCE 374 AA: 40645 MW: 74AB25950248F56F CRC64:
CC      -----
CC      Query Match: 52.4%; Score 118; DB 1; Length 374;
CC      Best Local Similarity 60.0%; Pred. No. 2,90e-08;
CC      Matches 18; Conservative 4; Mismatches 7; Indels 1; Gaps 1;
CC      -----
DB 92 IFSNNMNIKMPNYAGYKTFDGRCAQV 120
CC 1 IFSNNMNIKMPNYAGNKRRIKRVSNV 30
CC      -----
CC      RESULT 2
CC      YMV6 YEAST STANDARD: PRT: 898 AA.
CC      Q04748:
CC      01-NOV-1997 (Rel. 35, Created)
CC      01-NOV-1997 (Rel. 35, Last sequence update)
CC      01-NOV-1997 (Rel. 35, Last annotation update)
CC      HYPOTHEICAL 104.7 KD PROTEIN IN NCBI-HMS1 INTERGENIC REGION.
CC      YR0065W OR YN69.6.05.
CC      Saccharomyces cerevisiae (Baker's yeast).
CC      Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales
CC      Saccharomycetaceae; Saccharomyces.
CC      [1]
CC      SEQUENCE FROM N.A.
CC      STRAIN: S288C / AB972;
CC      Pearson D., Bowman S., Barrell B.G., Rejandream M.A., Walsh S.V.;
CC      Submitted (APR-1995) to the EMBL/GenBank/DBS databases.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: 248952; GAA88791.1;
CC      Hypothetical protein.
CC      SEQUENCE 898 AA: 104747 MW: 5707F3ACESC751F7 CRC64:
CC      -----
CC      Query Match: 35.9%; Score 83; DB 1; Length 898;
CC      Best Local Similarity 37.0%; Pred. No. 6.32e-02;
CC      Matches 10; Conservative 11; Mismatches 5; Indels 1; Gaps 1;
CC      -----
DB 785 DLNIDYDLSMYLKKLRQYLRLDISNI 811

```

```

QY 5 NLNIK_NMPLYIA-GNKRRIKRVSNV 30
CC      -----
CC      RESULT 3
CC      ID RM05_ACACA STANDARD: PRT: 177 AA.
CC      AC P46764;
CC      01-NOV-1995 (Rel. 32, Created)
CC      01-NOV-1995 (Rel. 32, Last sequence update)
CC      15-FEB-2000 (Rel. 39, Last annotation update)
CC      DE MITOCHONDRIAL 60S RIBOSOMAL PROTEIN L5.
CC      GN RPL5.
CC      OS Acanthamoeba castellanii (A. eba).
CC      OG Mitochondrion.
CC      OX Eukaryota; Acanthamoebidae; Entamoebae.
CC      RN [1]
CC      RP SEQUENCE FROM N.A.
CC      RC STRAIN: ATCC 30010 / NEFF;
CC      RX MEDLINE: 95147275.
CC      RA Burger G., Pante I., Loneragan K.M., Gray M.W.;
CC      R: "The mitochondrial DNA of the amoeboid protozoan, Acanthamoeba
CC      castellanii, complete sequence, gene content and gene
CC      organization."
CC      RL J. Mol. Biol. 245:522-537(1995).
CC      CC - SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC      CC - SIMILARITY: BELONGS TO THE LSP FAMILY OF RIBOSOMAL PROTEINS.
CC      CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL: U12386; AAD11844.1;
CC      DR PROSITE: PS00358; RIBOSOMAL_L5; FALSE_NEG.
CC      DR PFAM: PF00281; Ribosomal_L5.
CC      KW Ribosomal protein; Mitochondrion.
CC      SQ SEQUENCE 177 AA: 21764 MW: BEA193UDRC37E274 CRC64:
CC      -----
CC      Query Match: 36.0%; Score 81; DB 1; Length 177;
CC      Best Local Similarity 50.0%; Pred. No. 1.34e-01;
CC      Matches 13; Conservative 6; Mismatches 4; Indels 3; Gaps 3;
CC      -----
DB 46 RSLKFIKLYLMILYIS-NQKPFITKV 70
CC 4 KMLN-IKLN-MPDIANKKPFITKV 27
CC      -----
CC      RESULT 4
CC      ID PEL_TOBAC STANDARD: PRT: 397 AA.
CC      AC P40572;
CC      01-FEB-1995 (Rel. 31, Created)
CC      01-FEB-1995 (Rel. 31, Last sequence update)
CC      15-DEC-1998 (Rel. 37, Last annotation update)
CC      DE PECTATE LYASE PRECURSOR (EC 4.2.2.2).
CC      CS Nicotiana tabacum (Common tobacco).
CC      OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC      eudicotyledons; Magnoliophyta; eudicotyledons;
CC      core eudicots; Asteridae; euasterids I; Solanales; Solanaceae;
CC      Nicotiana.
CC      RN [1]
CC      RP SEQUENCE FROM N.A.
CC      RC STRAIN: CV. SAMSUN; TISSUE: POLLEN;
CC      RX MEDLINE: 93043039.
CC      RA Rogers H.J., Harvey A., Lonsdale D.M.;
CC      R: "Isolation and characterization of a tobacco gene with homology to
CC      pectate lyase which is specifically expressed during
CC      microsporogenesis."
CC      RL Plant Mol. Biol. 20:493-502(1992).
CC      CC - CATALYTIC ACTIVITY: ELIMINATIVE CLEAVAGE OF PECTATE TO GIVE
CC      OLIGOSACCHARIDES WITH 4-DEOXY-ALPHA-D-MANN-4-ENURONOSYL GROUPS

```

CC AT THEIR NON-REDUCING ENDS.
 CC -:- DEVELOPMENTAL STAGE: MAXIMALLY EXPRESSED LATE IN POLLEN
 CC DEVELOPMENT.
 CC -:- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X67158; CAA47630.1;
 CC EMBL: X67159; CAA47631.1;
 CC EMBL: X61102; CAA43414.1;
 CC PIR: S26211; S25211;
 CC PFAM: PF00544; pect_lyase_1;
 CC PRINTS: PR09007; AMBALLEPEN.
 CC Lysase: Signal.
 CC SIGNAL 1 25 POTENTIAL.
 CC CHAIN 26 397 PECTATE LYASE.
 CC ACT_SITE 272 272 POTENTIAL.
 CC CARBOHYD 134 134 POTENTIAL.
 CC CARBOHYD 227 227 POTENTIAL.
 CC CONFLICT 156 156 S -> C (IN MRNA).
 CC CONFLICT 189 190 GS -> SG (IN MRNA).
 CC CONFLICT 200 200 S -> D (IN MRNA).
 CC CONFLICT 202 202 H -> R (IN MRNA).
 CC CONFLICT 249 249 H -> N (IN MRNA).
 CC SEQUENCE 397 AA; 44351 MW; EF0A82CE5DA7543F CRC64;
 CC
 CC Query Match 36.0%; Score 81; DB 1; Length 397;
 CC Best Local Similarity 50.0%; Pred. No. 1; te-01;
 CC Matches 10; Conservative 7; Mismatches 3; Indels 0; Gaps 0;
 CC
 CC DB 116 IFGKMKIKLSRELIVTSNK 135
 CC 116:||||:||||:||||:
 CC 1 IFSKNLIKLMPLYIAGNK 20
 CC
 CC RESULT 5
 CC ID POLG_PVRA STANDARD: PRT: 3140 AA.
 CC AC P17767;
 CC DT 01-DEC-1990 (Rel. 15, Created)
 CC DT 15-FEB-1994 (Rel. 28, Last sequence update)
 CC DT 15-DEC-1998 (Rel. 37, Last annotation update)
 CC DE GENOME POLYMERASE (CONTAINS: N-TERMINAL PROTEIN (P1); HELPER
 CC COMPONENT PROTEINASE (EC 3.4.22.-) (HC-PRO); PROTEIN P3; 6 KD PROTEIN
 CC 2 (6K2); CYTOPLASMIC INCLUSION PROTEIN (CI); 6 KD PROTEIN 2 (6K2);
 CC GENOME-LINKED PROTEIN (VPG); NUCLEAR INCLUSION PROTEIN A (NI-A) (NIA)
 CC (EC 3.4.22.-) (49 KD PHOSPHATASE) (49 KD-PRO); NUCLEAR INCLUSION
 CC PROTEIN B (NI-B) (NIB) (RNA-DIRECTED RNA POLYMERASE) (EC 2.7.7.48);
 CC COAT PROTEIN (CP)).
 CC OS Plum pox potyvirus (strain Rakovic) (ppv).
 CC OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;
 CC CC potyvirus.
 CC [1]
 CC RN SEQUENCE FROM N.A.
 CC RX MEDLINE: 89370814.
 CC RA Lain S., Riechmann J.L., Garcia J.A.;
 CC ST "The complete nucleotide sequence of plum pox potyvirus RNA.";
 CC PL Virus Res. 13:157-172(1989).
 CC EN
 CC PP SEQUENCE OF 1778-2342 FROM N.A.
 CC RX MEDLINE: 89268456.
 CC RA Garcia J.A., Riechmann J.L., Lain S.;
 CC RT "Proteolytic activity of the plum pox potyvirus NIA-like protein in
 CC Escherichia coli.";
 CC PL Virology 170:362-369(1989).
 CC RN [3];
 CC RP SEQUENCE OF 2263-3140 FROM N.A.
 CC RA Lain S., Riechmann J.L., Mendez E., Garcia J.A.;

RT "Nucleotide sequence of the terminal region of plum pox potyvirus
 RNA."
 RL Virus Res. 10:325-342(1988).
 CC -:- FUNCTION: CYTOPLASMIC INCLUSION PROTEIN HAS HELICASE ACTIVITY IT
 CC MAY BE INVOLVED IN REPLICATION.
 CC -:- FUNCTION: NUCLEAR INCLUSION PROTEIN A HAS PROTEOLYTIC ACTIVITY.
 CC -:- PTM: VPG IS COVALENTLY LINKED TO THE GENOMIC RNA.
 CC -:- PTM: THE VIRAL RNA OF POTYVIRUSES IS EXPRESSED AS A SINGLE
 CC POLYPROTEIN WHICH UNDERGOES POSTTRANSLATIONAL PROTEOLYTIC
 CC PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT
 CC INDIVIDUAL PROTEINS.
 CC -:- SIMILARITY: HC PROTEINASES BELONGS TO PEPTIDASE FAMILY C6
 CC -:- SIMILARITY: NI-A PROTEINASES BELONGS TO PEPTIDASE FAMILY C4
 CC -:- SIMILARITY: NI-B PROTEINASES BELONGS TO POTYVIRUSES POLYPROTEIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M21847; AA85458.1;
 CC EMBL: M26965; AAA47085.1;
 CC PIR: A50009; GNVSR.
 CC PFAM: PF00270; DEAD.
 CC PFAM: PF00680; RNA_dep_RNA_pol_1;
 CC PFAM: PF00767; Poty_coat_1;
 CC PFAM: PF00851; Peptidase_C6_1;
 CC PFAM: PF00863; Peptidase_C4_1;
 CC PFAM: PF01577; Poty_P1_1;
 CC PRINTS: PR03966; NIAPOTYPIASE.
 CC Hydrolase: Transferase; Thiol protease; RNA-directed RNA polymerase;
 CC Coat protein; Polyprotein; Covalent protein-RNA linkage; Helicase.
 CC ATP-binding.
 CC CHAIN 1 301 N-TERMINAL PROTEIN.
 CC CHAIN 302 914 HELPER COMPONENT PROTEINASE.
 CC CHAIN 915 1168 6 KD PROTEIN 1.
 CC CHAIN 1169 1803 CYTOPLASMIC INCLUSION PROTEIN.
 CC CHAIN 1804 1856 6 KD PROTEIN 2.
 CC CHAIN 1857 2292 GENOME-LINKED PROTEIN.
 CC CHAIN 2293 2810 NUCLEAR INCLUSION PROTEIN A.
 CC CHAIN 2811 3140 NUCLEAR INCLUSION PROTEIN B.
 CC CHAIN 1919 COAT PROTEIN.
 CC BINDING 1919 COVALENT LINKAGE OF VIRAL RNA (BY
 CC SIMILARITY).
 CC NP_BIND 1253 1260 ATP (POTENTIAL).
 CC SEQUENCE 3140 AA; 355577 MW; EF30BB0794203F5 CRC64.
 CC
 CC Query Match 36.0%; Score 81; DB 1; Length 3140;
 CC Best Local Similarity 46.7%; Pred. No. 1; 34e-01;
 CC Matches 14; Conservative 8; Mismatches 5; Indels 3; Gaps 3;
 CC
 CC DB 2410 IFS-ALNKAAGVALLY-SGKKRDYFKNVSD 2437
 CC 1 IFSKNLIKLMPLYIAGNKKRFKRVSN 29
 CC
 CC RESULT 6
 CC ID POLG_PVRA STANDARD: PRT: 3141 AA.
 CC AC P13529; Q84929; P89038;
 CC DT 01-JAN-1990 (Rel. 13, Created)
 CC DT 15-FEB-2000 (Rel. 39, Last sequence update)
 CC DT 15-FEB-2000 (Rel. 39, Last annotation update)
 CC DE GENOME POLYMERASE (CONTAINS: N-TERMINAL PROTEIN (P1); HELPER
 CC COMPONENT PROTEINASE (EC 3.4.22.-) (HC-PRO); PROTEIN P3; 6 KD PROTEIN
 CC 1 (6K1); CYTOPLASMIC INCLUSION PROTEIN (CI); 6 KD PROTEIN 2 (6K2);
 CC GENOME-LINKED PROTEIN (VPG); NUCLEAR INCLUSION PROTEIN A (NI-A) (NIA)
 CC (EC 3.4.22.-) (49 KD PHOSPHATASE) (49 KD-PRO); NUCLEAR INCLUSION
 CC PROTEIN B (NI-B) (NIB) (RNA-DIRECTED RNA POLYMERASE) (EC 2.7.7.48);
 CC COAT PROTEIN (CP)).

US PLUM POX POTYVIRUS (strain D) (PPV).
 CC VIRUSES: ssRNA positive-strand viruses, no DNA stage, Potyviridae.
 CC POTYVIRUS.

PN SEQUENCE FROM N.A.
 RX MEDLINE: 90056790.
 RA Teycheney P.Y., Taverit G., Delbos R., Ravelonandri M., Dunez J.,
 RI "The complete nucleotide sequence of plum pox virus RNA (strain D)."
 RL Nucleic Acids Res. 17:1015-1016(1989).
 RN (2).
 RP SEQUENCE OF 2810-3141 FROM N.A., AND SEQUENCE OF 2812-2828.
 RA Ravelonandri M., Varveri C., Delbos R., Dunez J.,
 RI "Nucleotide sequence of the capsid protein gene of plum pox
 RT potyvirus".
 RL J. Gen. Virol. 69:1509-1516(1988).
 RN (3).
 RP REVISIONS TO C-TERMINUS.
 RA le Gal, G.:
 RC Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
 CC - FUNCTION: HELPER COMPONENT-PROTEINASE IS REQUIRED FOR APHID
 CC TRANSMISSION AND ALSO HAS PROTEOLYTIC ACTIVITY.
 CC - FUNCTION: CYTOPLASMIC INCLUSION PROTEIN HAS HELICASE ACTIVITY. IT
 CC MAY BE INVOLVED IN REPLICATION.
 CC - FUNCTION: NUCLEAR INCLUSION PROTEIN A HAS PROTEOLYTIC ACTIVITY.
 CC - P.N.: VPG IS COVALENTLY LINKED TO THE GENOMIC RNA.
 CC - P.N.: THE VIRAL RNA OF POTYVIRUSES IS EXPRESSED AS A SINGLE
 CC POLYPEPTIDE WHICH UNDERGOES POSTTRANSLATIONAL PROTEOLYTIC
 CC PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT
 CC INDIVIDUAL PROTEINS.
 CC - SIMILARITY: HC PROTEINASE BELONGS TO PEPTIDASE FAMILY C6.
 CC - SIMILARITY: NI-A PROTEINASE BELONGS TO PEPTIDASE FAMILY C4.
 CC - SIMILARITY: BELONGS TO THE POTYVIRUSES POLYPEPTIDE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed, used by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: X16415; CAA34437.1;
 DR EMBL: 000298; BAB002.0.1;
 DR FIC: S05929; GNVSFD.
 DR FIC: ZAO079; ZAO078.
 DR PFAM: PF00270; TRAC1.
 DR PFAM: PF00560; RNA_dep_PNA_P1.
 DR PFAM: PF00767; Potyvirus.
 DR PFAM: PF00451; Euplidasin.
 DR PFAM: PF03563; Euplidasin.
 DR PFAM: PF01577; Poly.
 DR PRINTS: PR00986; NACPPTASE.
 KW Hydroxylase; Transferase; Thiol protease; RNA-directed RNA polymerase;
 KW Coat protein; Polyprotein; Covalent protein-RNA linkage; Helicase;
 KW ATP-binding.

FT CHAIN 1 301 N-TERMINAL PROTEIN.
 FT CHAIN 322 915 HELPER COMPONENT PROTEINASE.
 FT CHAIN 316 2 6 KD PROTEIN A.
 FT CHAIN ? 1169 PROTEIN P3.
 FT CHAIN 1170 1804 CYTOPLASMIC INCLUSION PROTEIN.
 FT CHAIN 1205 1857 6 KD PROTEIN 2.
 FT CHAIN 1858 ? GENOME-LINKED PROTEIN.
 FT CHAIN ? 2293 NUCLEAR INCLUSION PROTEIN A.
 FT CHAIN 2294 2811 NUCLEAR INCLUSION PROTEIN B.
 FT CHAIN 2812 3141 COAT PROTEIN.
 FT BINDING 1920 1920 COVALENT LINKAGE OF VIRAL RNA (BY
 FT SIMILARITY).
 FT NP_S:NC 1254 1261 ATP (POTENTIAL).
 FT SEQUENCE 3141 AA: 355569 MW: 633641C404414DB3 CRC64:

Query Match 36.0% Score 81 DB 1: Length 3141:
 Res. Local Similarity 46.7% Pred. NO. 1.34e-01:
 Matches 14: Conservative 8: Mismatches 5: Indels 3: Gaps 3:

DB 2421 IFS-ALNKAAGALY-SOKKROYEKNVSD 2438
 QY 1 IFSKNLNTKLN-PLYAGNKKRF-KRVSN 29

RESULT 7
 ID YBR7 YEAS1 STANDARD: PRI: 143 AA
 AC P29183;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE VERY HYPOTHEICAL 14.3 KD PROTEIN IN NUP70-115; INTERMEDIATE BETWEEN
 GN YBL077W OR YBL0733.
 CS Saccharomyces cerevisiae (Baker's Yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales
 CC Saccharomycetaceae; Saccharomyces.
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C;
 RX MEDLINE: 96076535.
 RA Oberlander B., Gassenhuber J., Pittavandi E., Dandey H.,
 RT "Sequence analysis of a 78.6 kb segment of the left end of
 RT Saccharomyces cerevisiae chromosome III".
 RL Yeast 11:1103-1112(1995).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C;
 RA Contreras R., Fiers W., Ioshita M., Molenens F.,
 RL Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed, used by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: X79489; CAA56033.1;
 DR EMBL: 235838; CAA44897.1;
 DR PUR: S45433; S45433.
 KW Hypothetical protein.
 SO SEQUENCE 143 AA: 16327 MW: 10653FFC09794CCD CRC64:

Query Match 44.7% Score 78 DB 1: Length 143:
 Res. Local Similarity 33.9% Pred. NO. 4.1e-01:
 Matches 7: Conservative 12: Mismatches 2: Indels 0: Gaps 0:

DB 23 NUNVNHVYVLS-SRKKIYK 43
 QY 5 NUNIKAMPYIAGNKKRFK 25

RESULT 8
 ID YRR5 CAEEL STANDARD: PRI: 532 AA
 AC G39346;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE HYPOTHEICAL 59.6 KD PROTEIN R144.5 IN CHROMOSOME III.
 GN R144.5.
 CS Caenorhabditis elegans.
 CC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;
 CC Rhabditina; Rhabditoidea; Rhabditidae; Pelodidae; Caenorhabditis.
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2;
 RA Faveillo T.;
 RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL: U23515; AAC34545.1; ..
 CC WORMPEP: R144.5; CE02035.
 CC Hypothetical protein.
 CC SEQUENCE 532 AA: 59809 MW: 103588FF445B4283 CRC64:

Query Match 33.8% Score 76; DB 1; Length 532;
 Best Local Similarity 35.5% Pred. No. 8.38e-01;
 Matches 11; Conservative 11; Mismatches 7; Indels 2; Gaps 2;

DB 160 LFSANAK-KLVSPVLC-NPRLEIRHVTDIV 208
 QY 1 IFSKNLNKLNMPYAGNKRREIKRVSNVI 31

RESULT 9
 ID YA222-METCA STANDARD: PRT: 569 AA.
 AC Q58428:
 D1 15-JUL-1998 (Rel. 36, Created)
 D1 15-JUL-1998 (Rel. 36, Last sequence update)
 D1 15-JUL-1998 (Rel. 36, Last annotation update)
 DE HYPOTHETICAL PROTEIN M31022.
 GN M31022
 OS Methanococcus jannaschii
 CC Archaea: Euryarchaeota; Methanococcales; Methanococcaceae;
 CC Methanococcus
 CC [1]
 SEQUENCE FROM N.A.
 RC STRAIN=AL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE: 9631759;
 RA Bait C.J., White G., Olson G.J., Zhou L., Fleischmann R.D.,
 RA Sitten G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Goynne J.D.,
 RA Kariavade A.R., Dougherty B.A., Tob J.-F., Adams M.D., Reich C.L.,
 RA Overbeck R., Kirkness E.F., Weinstein K.G., Metrick J.M., Glodek A.,
 RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
 RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.,
 RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
 RA jannaschii".
 RC Science 273:1058-1073(1996).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: U67545; AAB93026.1; ..
 CC TIGR: M31022.
 CC FRAM: PF01472; PUA; 1.
 CC Hypothetical protein.
 CC SEQUENCE 569 AA: 55581 MW: E29EC27908E366ED CRC64:

Query Match 33.8% Score 76; DB 1; Length 569;
 Best Local Similarity 40.7% Pred. No. 8.38e-01;
 Matches 11; Conservative 8; Mismatches 6; Indels 2; Gaps 2;

DB 102 FSKHVDLIEDADLYIADGRLLIER 128
 QY 2 FSKNLN-1KL-NMPLYAGNKRREIKRV 26

RESULT 10
 ID RH31-YFAST STANDARD: PRT: 347 AA.
 AC Q06624:

DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE DNA DAMAGE TOLERANCE PROTEIN RHC31 (RAD31 HOMOLOG).
 GN AOS1 OR RHC3; OR YPR183W OR P97051.5.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetidae;
 CC Saccharomycetaceae; Saccharomycetes.
 CC [1]
 CC SEQUENCE FROM N.A.
 RC STRAIN=S288C / AB972;
 RA Johnston M., Andrews S., Brinkman M., Cooper J., Ding H., Du Z.,
 RA Kucaba T., Fulton L., Gattung S., Greco T., Kirsten J.,
 RA Kucaba T., Wallsworth K., Hawkins J., Hillier L., Jier M.,
 RA Johnson D., Johnston L., Lanston Y., Latreille P., Le T.,
 RA Mardis E., Menezes S., Miller N., Nhan M., Pauley A., Peluso D.,
 RA Rifkin L., Riles L., Raich A., Trevaskis E., Vignati D.,
 RA Wilcox L., Mohdram P., Vaudin M., Wilson R., Waterston R.,
 RL Submitted (MAY-1995; to the EMBL/GenBank/DBJ databases
 RN [2]
 RP CHARACTERIZATION:
 RX MEDLINE: 97248684.
 RA Shayeghi M., Does C.L., Tavassoli M., Watts F.Z.,
 RA "Characterisation of Schizosaccharomyces pombe rad31, a USA-related
 RA gene required for DNA damage tolerance".
 RL Nucleic Acids Res. 25:1162-1169(1997).

CC -1- FUNCTION: COULD BE INVOLVED IN A URICUIN-N-RELATED PROCESS
 CC IMPORTANT FOR DNA DAMAGE TOLERANCE.
 CC -2- SIMILARITY: TO THE N-TERMINAL OF URICUIN-ACTIVATING ENZYME E1.
 CC -3- SIMILARITY: STRONG, TO S.POMBE RAD31.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: U25842; AAB68113.1; ..
 CC SGD: L0004031; AOS1.
 CC FRAM: PF00899; Thif-family; 1.
 CC DNA damage. 347 AA: 39273 MW: 1AD7B817RA277F9E CRC64:
 CC SEQUENCE 347 AA: 39273 MW: 1AD7B817RA277F9E CRC64:

Query Match 33.3% Score 75; DB 1; Length 347;
 Best Local Similarity 50.0% Pred. No. 1.20e-02;
 Matches 10; Conservative 6; Mismatches 5; Indels 1; Gaps 0;

DB 140 NTLRKLNIPLYVAGSNLIF 159
 QY 4 KNLNKLNMPYIAGNKRRE 23

RESULT 11
 ID AX6B-SOYBN STANDARD: PRT: 90 AA.
 AC P33083:
 D1 01-OCT-1993 (Rel. 27, Created)
 D1 01-OCT-1993 (Rel. 27, Last sequence update)
 D1 01-OCT-1993 (Rel. 27, Last annotation update)
 DE AUXIN-INDUCED PROTEIN 6B.
 OS Glycine max (Soybean)
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 CC core eudicots; Rosidae; eurosoids 1; Fabales; Fabaceae; Papilionoideae;
 CC Glycine.
 CC [1]
 CC SEQUENCE FROM N.A.
 RC STRAIN=CV WAYNE;
 RX MEDLINE: 92404712.
 RA McClure B.A., Hagen G., Brown C.S., Gee M.A., Guilfoyle T.J.,
 RA "transcription, organization, and sequence of an auxin-regulated gene
 RA cluster in soybean".


```

CC-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/annouce/
CC or send an email to license@isb-sib.ch).
CC-----
CC EMBL: M35287; AAA42087.1;
CC EMBL: M35298; AAA42088.1;
CC PIR: A35639; A35639.
CC SCROB: GCR.0159;
CC PROSITE: P500237; G-PROTEIN_RECEPTOR; 1.
CC Pfam: PF00031; 7cml.1.
CC G-protein coupled receptor; Transmembrane: Glycoprotein.
CC DOMAIN 1 44 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 45 66 1 (POTENTIAL).
CC DOMAIN 67 82 2 (POTENTIAL).
CC TRANSMEM 83 104 2 (POTENTIAL).
CC DOMAIN 105 123 3 (POTENTIAL).
CC TRANSMEM 124 144 3 (POTENTIAL).
CC DOMAIN 145 160 4 (POTENTIAL).
CC TRANSMEM 161 181 4 (POTENTIAL).
CC DOMAIN 182 198 5 (POTENTIAL).
CC TRANSMEM 199 220 5 (POTENTIAL).
CC DOMAIN 221 241 6 (POTENTIAL).
CC TRANSMEM 242 263 6 (POTENTIAL).
CC DOMAIN 264 273 7 (POTENTIAL).
CC TRANSMEM 274 294 7 (POTENTIAL).
CC DOMAIN 295 343 7 (POTENTIAL).
CC CARBOHYD 4 4 POTENTIAL.
CC SEQUENCE 343 AA; 38364 MW; E4630007770941F4 CRC64;
CC-----
CC Query Match 32.4%; Score 73; DB 1; Length 343;
CC Best Local Similarity 32.3%; Pred. No. 2.42e+00;
CC Matches 10; Conservative 10; Mismatches 10; Indels 1; Gaps 1;
CC-----
DE 99 GSKAVIALNMGIFLGSFPDYVVRVRSIV 122
GY 1 FSKNENIKLNMPLYAGNRKRREIKRVSNV 31
CC-----
RESULT 15
CC POLG-PPVNA STANDARD: PRT: 3125 AA.
CC P1756;
CC 01-AUG-1990 (Rel. 15, Created)
CC 01-AUG-1990 (Rel. 15, Last sequence update)
CC 15-DEC-1999 (Rel. 37, Last annotation update)
CC GENOME POLYPEPTIDE [CONTAINS N-TERMINAL PROTEIN (P.)]; HELPER
CC COMPONENT PROTEINASE (EC 3.4.22.-) (HC-PRO); PROTEIN P.; 6 KD PROTEIN
CC 1 (6K1); CYTOPLASMIC INCLUSION PROTEIN (CI); 6 KD PROTEIN 2 (6K2);
CC GENOME-LINKED PROTEIN (VPG); NUCLEAR INCLUSION PROTEIN A (NI-A) (NIA)
CC (EC 3.4.22.-) (49 KD PROTEINASE) (49 KD-PRO); NUCLEAR INCLUSION
CC PROTEIN B (NI-B) (RNA-DIRECTED RNA POLYMERASE) (EC 2.7.7.48));
CC COAT PROTEIN (CP);
CC Plum pox polyvirus (isolate NAT) (PPV);
CC Viruses; ssRNA positive-strand viruses, no DNA stage: Potyviridae;
CC Potyviruses;
CC ;1)
CC SEQUENCE FROM N.A.
CC MEDLINE: 99279232.
CC RA Maiss E., Timpe J., Brisse A., Jekmann W., Casper R., Himmeler G.,
CC RA Matcovich D., Katinger H.W.D.;
CC "The complete nucleotide sequence of plum pox virus RNA.";
CC 1. J. Gen. Virol. 70:513-524(1989).
CC -1- FUNCTION: HELPER COMPONENT-PROTEINASE IS REQUIRED FOR APHID
CC TRANSMISSION AND ALSO HAS PROTEOLYTIC ACTIVITY.
CC -1- FUNCTION: CYTOPLASMIC INCLUSION PROTEIN HAS HELICASE ACTIVITY. IT
CC MAY BE INVOLVED IN REPLICATION
CC -1- FUNCTION: NUCLEAR INCLUSION PROTEIN A HAS PROTEOLYTIC ACTIVITY.
CC -1- PTM: VPG IS COVALENTLY LINKED TO THE GENOMIC RNA.
CC -1- PTM: THE VIRAL RNA OF POTYVIRUSES IS EXPRESSED AS A SINGLE

```

```

CC POLYPROTEIN WHICH UNDERGOES POSTTRANSLATIONAL PROTEOLYTIC
CC PROCESSING RESULTING IN THE PRODUCTION OF AT LEAST EIGHT
CC INDIVIDUAL PROTEINS.
CC
CC -1- SIMILARITY: HC PROTEINASE BELONGS TO PEPTIDASE FAMILY C6.
CC -1- SIMILARITY: N-A PROTEINASE BELONGS TO PEPTIDASE FAMILY C4.
CC -1- SIMILARITY: BELONGS TO THE POLYVIRUSES POLYPROTEIN FAMILY.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by not-for-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to licens@isb-sib.ch).
CC
CC EMBL: D13751; BAA02898.1;
CC PIR: JQ0003; GNVSP.
CC
CC PFAM: PF00270; DEAD; 1.
CC
CC PFAM: PF00680; RNA_dep_RNA_pol; 1.
CC
CC PFAM: PF00757; Poty_coat; 1.
CC
CC PFAM: PF00851; Peptidase_C6; 1.
CC
CC PFAM: PF00863; Peptidase_C4; 1.
CC
CC PFAM: PF01577; Poty_P1; 1.
CC
CC PRINTS: PRO0966; NIAPOTYPEASE.
CC
CC Hydrolase; Transferase; Thiol protease; RNA-directed RNA polymerase;
CC Coat protein; Polyprotein; Covalent protein-RNA linkage; Helicase
CC ATP-binding.
CC
CC CHAIN 1 301 N-TERMINAL PROTEIN.
CC CHAIN 302 914 16 PER COMPONENT PROTEINASE.
CC CHAIN 915 ? 1; TEIN P3.
CC
CC CHAIN ? 1168 6 KD PROTEIN 1.
CC CHAIN 1169 1803 CYTOPLASMIC INCLUSION PROTEIN
CC CHAIN 1804 1856 6 KD PROTEIN 2.
CC
CC CHAIN 1857 ? GENOME-LINKED PROTEIN
CC CHAIN ? 2292 NUCLEAR INCLUSION PROTEIN A.
CC CHAIN 2293 2810 NUCLEAR INCLUSION PROTEIN H.
CC CHAIN 2811 3125 COAT PROTEIN.
CC BINDING 1919 1919 COVALENT LINKAGE OF VIRAL RNA (RY
CC SIMILARITY).
CC
CC NLRIND 1253 1260 ATP (POENTIAL).
CC SEQUENCE 3125 AA: 354261 MW: 800533439CB712 CRC64:
CC
CC Query Match 32.48; Score 73; DB 1; Length 3125;
CC Best Local Similarity 43.38; Pred. No. 2.42e-00;
CC Matches 13; Conservative 8; Mismatches 6; Indels 3; Gaps 3;
CC
CC Db 2410 IFS-ALNMAAAGALY-SKKKRYEVNVD 2417
CC
CC QY 1 IFSKNINIKLM-PLYIAGNKKRPIKPVN 29
CC
CC Search completed: Mon Jun 19 15:59:38 2000
CC Job time : 20 secs.

```

Query Match

Query Match 32.48; Score 73; DB 1; Length 3125;

Best Local Similarity 43.3%; Pred. No. 2.42e-00;
13; Matches 6; Mismatches 6; Indels 3; Gaps 3;

2410 -FS-A:NYKAAVGALY-S3KKKDYFENVSD 2417

[illegible]

Search completed: Mon Jun 19 15:59:38 2000
Job time : 20 secs.

Job time : 20 secs.

WATERBURY

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit,
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsearch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Mon Jun 19 15:59:55 2000: Master time 8.99 seconds
239.116 Million cell updates/sec
Tabular output not generated.

Title: >US-09-142-524A-4
Description: (1-31) from US09142524A.pep
Perfect Score: 225
Sequence: 1 IFSKNLNKLNMPLYAGNKRRTIKRVSNV 31

Scoring table: PAM 250
Gap 11

Searched: 225878 seqs. 69334.22 residues

Post-processing: Minimum Match 08
Listing first 45 summaries

Database: spiremb112

1:sp_archaea 2:sp_bacteria 3:sp_fungi 4:sp_human
5:sp_invertebrate 6:sp_mammal 7:sp_mhc 8:sp_organella
9:sp_phase 10:sp_plant 11:sp_rodent 12:sp_unclassified
13:sp_vertebrate 14:sp_virus

Statistics: Mean 33.518: Variance 57.768: scale 0.580

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description | Pred. No. |
|------------|-------|-------------|--------|-------|------------------------|-----------|
| 1 | 143 | 63.6 | 375 | 10 | CHAO1 PRECURSOR. | 2.83e-12 |
| 2 | 123 | 54.7 | 367 | 10 | POLLEN MAJOR ALLERGEN | 2.01e-08 |
| 3 | 199 | 39.6 | 707 | 5 | EO3H4.4 PROTEIN. | 2.17e-02 |
| 4 | 84 | 37.3 | 741 | 10 | ETHYLENE RECEPTOR | 1.40e-01 |
| 5 | 83 | 36.9 | 365 | 13 | SEQUENCE OF BAC F19K23 | 2.02e-01 |
| 6 | 81 | 36.0 | 480 | 2 | HYPOTHETICAL 54.0 KD P | 4.17e-01 |
| 7 | 78 | 34.7 | 740 | 10 | ETHYLENE RECEPTOR HOMO | 1.21e-00 |
| 8 | 77 | 34.2 | 73 | 5 | HYPOTHETICAL 8.6 KD PF | 1.72e-00 |
| 9 | 77 | 34.2 | 280 | 2 | ORF2 (FRAGMENT) | 1.72e-00 |
| 10 | 77 | 34.2 | 480 | 2 | HYPOTHETICAL 54.0 KD P | 1.72e-00 |
| 11 | 76 | 33.8 | 741 | 10 | ETHYLENE RECEPTOR HOMO | 2.44e-00 |
| 12 | 76 | 33.8 | 3140 | 14 | COAT PROTEIN (FRAGMENT | 2.44e-00 |
| 13 | 75 | 33.3 | 273 | 1 | HYPOTHETICAL PROTEIN M | 3.44e-00 |
| 14 | 75 | 33.3 | 1028 | 11 | NB-3. | 3.44e-00 |
| 15 | 74 | 32.9 | 331 | 2 | HYPOTHETICAL 37.6 KD P | 4.84e-00 |
| 16 | 74 | 32.9 | 637 | 5 | P7056K PROTEIN KINASE | 4.84e-00 |
| 17 | 74 | 32.9 | 637 | 5 | 70 KDA S6 KINASE. | 4.84e-00 |
| 18 | 72 | 32.0 | 133 | 10 | SEVEN IN ABSENTIA HOMO | 9.49e+00 |
| 19 | 72 | 32.0 | 331 | 14 | ACMNPV GRI13. | 9.49e+00 |
| 20 | 72 | 32.0 | 476 | 2 | UNDECAPRENYL-PHOSPHATE | 9.49e+00 |

| | | | | | | | |
|----|----|------|------|----|--------|------------------------|----------|
| 21 | 72 | 32.0 | 511 | 2 | Q9XDK3 | POTATIVE FLIPPASE. | 9.49e+00 |
| 22 | 72 | 32.0 | 586 | 10 | Q43390 | RCH2 PROTEIN. | 9.49e+00 |
| 23 | 72 | 32.0 | 689 | 14 | P83565 | 100 KDA PROTEIN. | 9.49e+00 |
| 24 | 72 | 32.0 | 963 | 2 | CC6458 | TREHALOSE SYNTHASE. | 9.49e+00 |
| 25 | 72 | 32.0 | 4981 | 5 | Q07172 | PFC0820W PROTEIN. | 9.49e+00 |
| 26 | 71 | 31.6 | 286 | 5 | Q1153 | SYNTAXIN. | 1.84e-01 |
| 27 | 71 | 31.6 | 288 | 2 | Q94RH8 | POTATIVE DIHYDROXYAC | 1.84e-01 |
| 28 | 71 | 31.6 | 331 | 5 | Q1829 | F15A4 7 PROTEIN. | 1.84e-01 |
| 29 | 71 | 31.6 | 331 | 5 | Q1829 | ETHYLENE RECEPTOR. | 1.84e-01 |
| 30 | 71 | 31.6 | 845 | 2 | Q948A6 | CLIPC PROTEASE. | 1.84e-01 |
| 31 | 71 | 31.6 | 854 | 2 | Q84289 | CLIPC PROTEASE ATPASE. | 1.84e-01 |
| 32 | 71 | 31.6 | 925 | 2 | Q48449 | TIG START CODON. | 1.84e-01 |
| 33 | 71 | 31.6 | 1039 | 13 | Q92250 | CONTACTIN A. | 1.84e-01 |
| 34 | 70 | 31.1 | 152 | 10 | Q9XIK5 | T10024.6. | 1.84e-01 |
| 35 | 70 | 31.1 | 240 | 14 | Q9WMP8 | VE10. | 1.84e-01 |
| 36 | 70 | 31.1 | 309 | 5 | Q62129 | F02H5 2 PROTEIN | 1.84e-01 |
| 37 | 70 | 31.1 | 472 | 2 | Q66533 | HYPOTHETICAL 55.9 KD P | 1.84e-01 |
| 38 | 70 | 31.1 | 613 | 10 | Q49431 | ETHYLENE RESPONSE SENS | 1.84e-01 |
| 39 | 70 | 31.1 | 613 | 10 | Q38446 | ETHYLENE RESPONSE SENS | 1.84e-01 |
| 40 | 70 | 31.1 | 618 | 2 | Q64435 | HYPOTHETICAL 71.2 KD P | 1.84e-01 |
| 41 | 70 | 31.1 | 760 | 10 | Q4795 | F2401.3. | 1.84e-01 |
| 42 | 70 | 31.1 | 796 | 5 | Q24592 | E239 9-1 (FRAGMENT). | 1.84e-01 |
| 43 | 70 | 31.1 | 1019 | 10 | P94378 | ALPHA-MANNOSIDASE. | 1.84e-01 |
| 44 | 70 | 31.1 | 1533 | 2 | Q84926 | HYPOTHETICAL 172.9 KD | 1.84e-01 |
| 45 | 69 | 30.7 | 1494 | 2 | C11381 | SENSORY TRANSDUCTION H | 2.55e-01 |

AL NMENTS

| | | | | |
|---|--|-----------------------------------|------------|---------|
| RESULT | 1 | PRELIMINARY: | PRT: | 375 AA. |
| ID | Q96385 | | | |
| AC | Q96385 | | | |
| DI | 01-FEB-1997 (TRENBLREL. 02. Created) | | | |
| DI | 01-FEB-1997 (TRENBLREL. 02. Last sequence update) | | | |
| DI | 01-NOV-1999 (TRENBLREL. 12. Last annotation update) | | | |
| DE | CHAO1 PRECURSOR | | | |
| OS | Chamaecyparis obtusa. | | | |
| OC | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; | | | |
| OC | euphyllophytes; Spermatophyta; Coniferopsida; Coniferales; | | | |
| OC | Taxodiaceae; Chamaecyparis. | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | TISSUE-POLLEN: | | | |
| RX | MEDLINE: 96265594. | | | |
| RA | SUZUKI M., KOMIYAMA N., ITOH M., ITOH H., SONE T., KURO K., TAKAGI I., | | | |
| RA | OHYA N. | | | |
| RT | Participation, characterization and molecular cloning of Chao1, a | | | |
| RL | major allergen of Chamaecyparis obtusa (Japanese Cypress) pollen. | | | |
| DR | EXBL: Q46434; BAA08247.1 | | | |
| DR | REDEL: 7626; Chao1:88.7474 | | | |
| DR | PFAM: PF00544; pecllyase:1. | | | |
| DR | PRINTS: PR00807; AMBALLENGEN. | | | |
| KW | Signal. | | | |
| FT | SIGNAL | 1 21 | POTENTIAL. | |
| FT | CHAIN | 22 375 | CHAO1 | |
| SQ | SEQUENCE | 375 AA: 40258 MW: A0961492 CRC32: | | |
| Query Match 63.6% Score 143: DB 10: Length 375; | | | | |
| Best Local Similarity 100.0% Pred. No. 2.83e-12; | | | | |
| Matches 20: Conservative 0: Mismatches 0: Indels 0: Gaps 0; | | | | |
| DB | 92 | IFSKNLNKLNMPLYAGNK 111 | | |
| QY | 2 | IFSKNLNKLNMPLYAGNK 20 | | |
| RESULT | 2 | PRELIMINARY: | PRT: | 367 AA. |
| ID | Q92NU7 | | | |
| AC | Q92NU7 | | | |
| DI | 01-MAY-1999 (TRENBLREL. 10. Created) | | | |
| DI | 01-MAY-1999 (TRENBLREL. 10. Last sequence update) | | | |
| DI | 01-NOV-1999 (TRENBLREL. 12. Last annotation update) | | | |

GN B9C24:
 OC Borrelia burgdorferi (Lyme disease spirochete).
 OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 35210 / B31;
 RX MEDLINE: 98065943
 RA FRASER C.M., CASSENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,
 RA LATHIGRA R., WHITE O., KETCHUM K.A., DODSON R., HICKEY E.K., GWINN M.,
 RA DOUGHERTY B., TOMB J.F., FEISCHMAN R.D., RICHARDSON D.,
 RA PETERSON J., KERLAVAGE A.R., QUACKENBUSH J., SALZBERG S., HANSON M.,
 RA VAN VEGT J., PALMER N., ADAMS M.D., GOCAYNE J.D., WEIDMAN J.,
 RA UTTERBACK T., MATTHEY L., McDONALD L., ARTIACH P., BOWMAN C.,
 RA GARLAND S., FUJII C., COTTON M.D., HORST K., ROBERTS K., HAICH B.,
 RA SMITH H.C., VENTER J.C.,
 RI "Genomic sequence of a Lyme disease spirochete, Borrelia
 RI burgdorferi."
 RL Nature 390:580-585(1997).
 DL EMBL: AE001137; AAC66591.1;
 DL TIGR: BB0261;
 KW Hypothetical protein.
 SQ SEQUENCE 460 AA: 54024 MW: 54143655 CRC32:
 Query Match 36.0%; Score 81; DB 2; Length 460;
 Best Local Similarity 33.3%; Pred. No. 4,17e-00;
 Matches 10; Conservative 10; Mismatches 10; Indels 0; Gaps 0;
 DB 320 FSKKNENIALRIYLRSKSEKFEKIANEI 349
 QY 2 FSKNNINIKNMPYIAGNKKRFKRVSNVI 31
 RESULT 7
 ID Q9XH58 PRELIMINARY: PRI: 740 AA.
 AC Q9XH58;
 DI 01-NOV-1999 (TREMBlrel. 12, Created)
 DI 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DI 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
 DE ETHYLENE RECEPTOR HOMOLOG.
 GN PRETI.
 OS Pelargonium hortorum.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eud; orydedons;
 OC core eudicots; Rosidae; Geraniales; Geraniaceae; Pelargonium.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC DERVINIS C., CLARK D.G.;
 RI "Effect of Pollination and Exogenous Ethylene on Accumulation of ETR;
 RI Homolog Transcripts During Flower Petal Abscission of Petunia
 RI (Pelargonium x hortorum L'H. Bal. var.)."
 RI Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 DL EMBL: AF141928; RAD3576.1;
 KW Receptor.
 SQ SEQUENCE 740 AA: 82956 MW: 96130075 CRC32:
 Query Match 34.7%; Score 78; DB 10; Length 740;
 Best Local Similarity 30.8%; Pred. No. 1,21e-00;
 Matches 8; Conservative 10; Mismatches 6; Indels 0; Gaps 0;
 DB 443 LNVSPDLPEVIGDEKRLVQIMLVV 468
 QY 6 LNIKLNMPYIAGNKKRFKRVSNVI 31
 RESULT 8
 ID Q46181 PRELIMINARY: PRI: 73 AA.
 AC Q46181;
 DI 01-JUN-1998 (TREMBlrel. 06, Created)
 DI 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
 DI 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
 DE HYPOTHETICAL 8.5 KD PROTEIN (FRAGMENT).
 OS Drosophila subobscura (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H211;
 RA BLESA D., MARTINEZ-SEBASTIAN M.J.;
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
 DL EMBL: U73032; AAB92390.1;
 DL FLYBASE: FBgn0023240; DsubNanon1;
 KW Hypothetical protein.
 FT NON-TER 1
 SQ SEQUENCE 73 AA: 8614 MW: 1F41786E CRC32:
 Query Match 34.2%; Score 77; DB 5; Length 73;
 Best Local Similarity 28.6%; Pred. No. 1,72e-00;
 Matches 8; Conservative 12; Mismatches 8; Indels 0; Gaps 0;
 DB 38 KKLNQVDMSLFFAGALSFSVLSINSIF 65
 QY 4 KNLNKLKMPYIAGNKKRFKRVSNVI 31
 RESULT 9
 ID Q9XE03 PRELIMINARY: PRI: 256 AA.
 AC Q9XE03;
 DI 01-NOV-1999 (TREMBlrel. 12, Created)
 DI 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DI 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
 DE ORF2 (FRAGMENT).
 OS Staphylococcus aureus.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;
 OC Bacillus/Staphylococcus group; Staphylococcus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NCIC8325;
 RX MEDLINE: 99278010.
 RA ITO T., KATAYAMA Y., HIRAMATSU K.;
 RI "Cloning and nucleotide sequence determination of the entire and DNA
 RI of pre-methicillin-resistant Staphylococcus aureus N315."
 RL Antimicrob. Agents Chemother. 43:1449-1458(1999).
 DL EMBL: AB014440; BAA82253.1;
 DL NON-TER 256
 FT NON-TER 256
 SQ SEQUENCE 256 AA: 30603 MW: F1FABCB3 CRC32:
 Query Match 34.2%; Score 77; DB 2; Length 256;
 Best Local Similarity 29.0%; Pred. No. 1,72e-00;
 Matches 9; Conservative 10; Mismatches 10; Indels 0; Gaps 0;
 DB 144 LFSKMGESLSIRVFSPKSLFKSLV 414
 QY 1 LFSKNNIKLNMPYIAGNKKRFKRVSNVI 31
 RESULT 10
 ID G65965 PRELIMINARY: PRI: 480 AA.
 AC G65965;
 DI 01-AUG-1998 (TREMBlrel. 07, Created)
 DI 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
 DI 01-NOV-1998 (TREMBlrel. 38, Last annotation update)
 DE HYPOTHETICAL 56.0 KD PROTEIN.
 GN AQ.765.
 OS Aquifex aeolicus.
 OC Bacteria; Aquificales; Aquificaceae; Aquifex.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VF5;
 RX MEDLINE: 98196666.
 RA DECKER G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
 RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,
 RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
 RI "The complete genome of the hyperthermophilic bacterium Aquifex
 RI aeolicus."
 RI Nature 392:353-358(1998).

Search completed: Mon Jun 19 16:00:15 2000
Job time : 20 secs.

```

RESULT 14
IC P97528 PRELIMINARY: PRI: 1028 AA.
AC P97528
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DI 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DI 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE NB-3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN 14.
RP SEQUENCE FROM N.A.
RC STRAIN=WI51A3; TISSUE=BRAIN;
RX MEDLINE: 97101230.
RA OGAWA J., KANEKO H., MASUDA T., NAGATA S., HOSoya H., KATANABE K.:
RT "Novel neural adhesion molecules in the Contactin/F3 subgroup of the
RT immunoglobulin superfamily: isolation and characterization of cDNAs
RT from rat brain." J. Biol. Chem. 270:173-176(1995).
RL Neurosci. Lett. 218:173-176(1996).
DR EMBL: D87248; BAA13320.1; -.
DR HSSP: P20241; ICFB.
DR PFAM: PF00041; fn3; 4.
DR PFAM: PF00347; ig; 5.
SC SEQUENCE 1028 AA; 114065 MW; 4798F057 CRC32;

Query Match 33.3%; Score 75; DB 11; Length 1028;
Best Local Similarity 58.3%; Pred. No. 3.44e+00;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

DL 163 PLYVQEDKRFV 174
QY 11111111
13 PLYAGNKRRI 24

RESULT 15
IC P74570 PRELIMINARY: PRI: 33; AA.
AC P74570
DT 02-FEB-1997 (TrEMBLrel. 02, Created)
DI 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DI 01-JAN-1995 (TrEMBLrel. 09, Last annotation update)
DE HYPOTHETICAL 37.6 KD PROTEIN.
OS Synechocystis sp. (Strain PCC 6803)
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
RN 11.
RP SEQUENCE FROM N.A.
RC STRAIN=PCC6803;
RX TABATA S.:
RL Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
RN 12.
RP SEQUENCE FROM N.A.
RC STRAIN=PCC6803;
RX MEDLINE: 97091201.
RA KANEKO T., SATO S., KOTANI H., TANAKA A., ASAMIZU E., NAKAMURA Y.,
RA MIYAJIMA N., HIRCSAWA M., SUGIURA M., SASAMOTO S., KIMURA T.,
RA HOSOGUCHI T., MATSUO A., MURAKI A., NAKAZAKI N., NARDO K., OKUMURA S.,
RA SHIMO S., TAKEUCHI C., WADA T., WATANABE A., YANADA M., YASUDA M.,
RA TABATA S.:
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions." J.
RL DNA Res. 3:109-136(1995).
DR EMBL: D90916; BAA18677.1; -.
KW Hypothetical protein.
SC SEQUENCE 341 AA; 37593 MW; B50A856C CRC32;

Query Match 32.9%; Score 74; DB 2; Length 331;
Best Local Similarity 32.0%; Pred. No. 4.84e+00;
Matches 8; Conservative 12; Mismatches 5; Indels 0; Gaps 0;

DL 4 NYQFEIALSFAGEDREYVDVRL 28
QY 11111111111111111111
7 NIKLNPLYIAGNKRRIKRVSNVI 31

```

WATERMAN

(TM)

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright: (c) 1994-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd
Waterman_pp protein - protein database search using Smith-Waterman algorithm
Run On: Mon Jun 19 16:03:25 2000: MasPar time 4.36 Seconds
Tabular output was generated. 168,264 Million cell updates/sec

Title: US-09-142-524A-5
Description: (1-31) from US09-142524A pep
Perfect Score: 207
Sequence: 1 SSGKNEGNIYNNNEAFKVERRFKRVSNVI 31
Scoring table: PAM 150
Gap 1:
Searched: 189463 seqs, 23686106 residues
Post-processing: Minimum Match 0%
Listing first 45 summaries
Database: a-geneseq35
l-genesep
Statistics: Mean 22.687; Variance 86.099; scale 0.264

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

| SUMMARIES | | | | | Pred. No. | |
|-----------|-----|-------|-------------|----------------------------------|-----------|--|
| Result | No. | Score | Query Match | Length DB ID Description | | |
| 1 | 207 | 109.5 | 31 | 1 W27373 Multi-epitope peptide | 1.63e-12 | |
| 2 | 225 | 60.4 | 20 | 1 W42154 T-cell epitope peptide | 3.41e-04 | |
| 3 | 125 | 60.4 | 354 | 1 W01344 Chamaecyparis obtusa p | 3.41e-04 | |
| 4 | 125 | 60.4 | 354 | 1 W42121 Japanese cypress polie | 3.41e-04 | |
| 5 | 125 | 60.4 | 375 | 1 W01345 Chamaecyparis obtusa p | 3.41e-04 | |
| 6 | 138 | 52.2 | 367 | 1 R45577 Jun s | 1.49e-02 | |
| 7 | 93 | 44.9 | 20 | 1 R82522 Cry j I Japanese Cedar | 3.80e-01 | |
| 8 | 93 | 44.9 | 20 | 1 R45573 Cry j I pollen allergen | 3.80e-01 | |
| 9 | 93 | 44.9 | 20 | 1 W83058 Control peptide P16. | 3.80e-01 | |
| 10 | 93 | 44.9 | 22 | 1 W80255 Sugil allergen protein | 3.80e-01 | |
| 11 | 93 | 44.9 | 26 | 1 R45599 Cry j I pollen allergen | 3.80e-01 | |
| 12 | 93 | 44.9 | 28 | 1 R92180 Cry j I Japanese Cedar | 3.80e-01 | |
| 13 | 93 | 44.9 | 30 | 1 R45598 Cry j I pollen allergen | 3.80e-01 | |
| 14 | 93 | 44.9 | 353 | 1 R75362 Japanese cedar pollen | 3.80e-01 | |
| 15 | 93 | 44.9 | 353 | 1 R81587 Cedar pollen allergen | 3.80e-01 | |
| 16 | 93 | 44.9 | 374 | 1 R31937 Cry j I. | 3.80e-01 | |
| 17 | 93 | 44.9 | 374 | 1 R60165 Japanese cedar pollen | 3.80e-01 | |
| 18 | 93 | 44.9 | 374 | 1 R45541 Cry j I pollen allergen | 3.80e-01 | |
| 19 | 93 | 44.9 | 374 | 1 R82490 Cry j I Japanese Cedar | 3.80e-01 | |
| 20 | 90 | 43.5 | 81 | 1 W80358 Sugil allergen protein | 7.17e-01 | |
| 21 | 99 | 43.0 | 19 | 1 Y01802 Peptide used for the t | 8.96e-01 | |
| 22 | 87 | 42.0 | 19 | 1 Y01801 Peptide used for the t | 1.35e-00 | |
| 23 | 85 | 41.1 | 18 | 1 Y01804 Peptide used for the t | 2.04e-00 | |

| ID | W27373 standard; peptide: 31 AA | 18 | 1 | Y01803 | Peptide used for the t | 3.09e-00 |
|----|---|-----|---|--------|-------------------------|----------|
| AC | W27373 | 31 | 1 | W27372 | Multi-epitope peptide | 3.80e-00 |
| DI | 24-MAR-1998 (first entry) | 24 | 1 | R45600 | Cry j I pollen allergen | 5.73e-00 |
| DE | Multi-epitope peptide used as immunotherapeutic agent #5. | 26 | 1 | R45601 | Cry j I pollen allergen | 5.73e-00 |
| KW | Multi-epitope peptide; immunotherapeutic agent; allergic diseases | 80 | 1 | W27369 | Multi-epitope peptide | 1.29e-01 |
| KW | T-cell epitope region; allergen; lymphocyte; immunoglobulin E. | 105 | 1 | W27370 | Multi-epitope peptide | 1.29e-01 |
| OS | Synthetic. | 36 | 1 | W27371 | Multi-epitope peptide | 1.29e-01 |
| PN | W09732600-A1. | 134 | 1 | W27372 | Multi-epitope peptide | 1.29e-01 |
| PD | 12-SEP-1997 | 34 | 1 | R04895 | Penicillinase-Inhibitor | 4.51e-01 |
| PR | 10-MAR-1997: J00740 | 70 | 1 | W45866 | Thermotoga OC1/4V endo | 4.28e-01 |
| PA | (MEIP) MEIJI MILK PROD CO Ltd. | 30 | 1 | W45870 | Thermotoga OC1/4V endo | 4.28e-01 |
| PI | Dairiki K, Iwama A, Kino K, Kame A, Sone T. | 69 | 1 | W59334 | Haemophilus influenzae | 5.20e-01 |
| DR | WPI: 97-470495/43. | 50 | 1 | W42155 | T-cell epitope peptide | 6.33e-01 |
| PZ | Peptide immunotherapeutic agent to treat allergic diseases - | 68 | 1 | Y00910 | S. aureus ffb protein | 6.33e-01 |
| PZ | consals; multi-epitope peptide containing T cell epitope regions | 32 | 1 | R4585 | Cedar pollen allergen | 7.69e-01 |
| PS | Claim 9; Page 32: 58pp; Japanese. | 32 | 1 | Y01805 | Peptide used for the t | 7.69e-01 |
| CC | The present sequence represents a multi-epitope peptide which is used as | 40 | 1 | W15543 | 50K-cellulase from Mel | 9.33e-01 |
| CC | a new immunotherapeutic agent. It comprises 1 cell epitope regions from 2 | 66 | 1 | R13490 | Human C4 binding prote | 9.33e-01 |
| CC | or more different allergens (preferably linked via arginine or lysine | 41 | 1 | R62504 | Large polypeptide sequ | 9.33e-01 |
| CC | directs), where the T cell epitope regions have a positivity index | 66 | 1 | R04166 | Nematode excretory pro | 1.13e-02 |
| CC | greater than 100 as measured in a patient group responding to the | 65 | 1 | W37735 | Cytochrome P450RA1 iso | 1.13e-02 |
| CC | allergen; have at least 70% reactivity with lymphocytes from patients | 65 | 1 | W48721 | Human cytoplasmic iso | 1.13e-02 |
| CC | responding to the allergen; and are not reactive with immunoglobulin E | | | | | |
| CC | (IgE) antibodies from patients responsive to the allergen. The agent can | | | | | |
| CC | be used to prevent and treat a wide variety of allergic diseases, e.g. by | | | | | |
| CC | desensitisation. Side effects, e.g. those mediated by IgE, are reduced. | | | | | |
| SC | Sequence 31 AA. | | | | | |

ALIGNMENTS

| RESULT | 1 | | | | | |
|---|---|--|--|--|--|--|
| ID | W27373 standard; peptide: 31 AA | | | | | |
| AC | W27373 | | | | | |
| DI | 24-MAR-1998 (first entry) | | | | | |
| DE | Multi-epitope peptide used as immunotherapeutic agent #5. | | | | | |
| KW | Multi-epitope peptide; immunotherapeutic agent; allergic diseases | | | | | |
| KW | T-cell epitope region; allergen; lymphocyte; immunoglobulin E. | | | | | |
| OS | Synthetic. | | | | | |
| PN | W09732600-A1. | | | | | |
| PD | 12-SEP-1997 | | | | | |
| PR | 10-MAR-1997: J00740 | | | | | |
| PA | (MEIP) MEIJI MILK PROD CO Ltd. | | | | | |
| PI | Dairiki K, Iwama A, Kino K, Kame A, Sone T. | | | | | |
| DR | WPI: 97-470495/43. | | | | | |
| PZ | Peptide immunotherapeutic agent to treat allergic diseases - | | | | | |
| PZ | consals; multi-epitope peptide containing T cell epitope regions | | | | | |
| PS | Claim 9; Page 32: 58pp; Japanese. | | | | | |
| CC | The present sequence represents a multi-epitope peptide which is used as | | | | | |
| CC | a new immunotherapeutic agent. It comprises 1 cell epitope regions from 2 | | | | | |
| CC | or more different allergens (preferably linked via arginine or lysine | | | | | |
| CC | directs), where the T cell epitope regions have a positivity index | | | | | |
| CC | greater than 100 as measured in a patient group responding to the | | | | | |
| CC | allergen; have at least 70% reactivity with lymphocytes from patients | | | | | |
| CC | responding to the allergen; and are not reactive with immunoglobulin E | | | | | |
| CC | (IgE) antibodies from patients responsive to the allergen. The agent can | | | | | |
| CC | be used to prevent and treat a wide variety of allergic diseases, e.g. by | | | | | |
| CC | desensitisation. Side effects, e.g. those mediated by IgE, are reduced. | | | | | |
| SC | Sequence 31 AA. | | | | | |
| Query Match 100.0%; Score 207; DB 1; Length 31; | | | | | | |
| Best Local Similarity 100.0%; Pred. No. 1.63e-12; | | | | | | |
| Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | | | | |
| DB | 1 SSGKNEGNIYNNNEAFKVERRFKRVSNVI: 31 | | | | | |
| QY | 1 SSGKNEGNIYNNNEAFKVERRFKRVSNVI: 31 | | | | | |
| RESULT 2 | | | | | | |
| SC | W42154 standard; peptide: 20 AA. | | | | | |
| AC | W42154 | | | | | |
| DI | 16-JUN-1998 (first entry) | | | | | |
| DE | T-cell epitope peptide 34 from Japanese cypress pollen antigen (Chao). | | | | | |
| KW | Japanese cypress pollen; antigen: T-cell epitope; Chao; Chao2; | | | | | |
| KW | diagnosis: allergy; spring tree pollen disease; pollinosis. | | | | | |

```

KW Chamaecyparis obtusa.
OS WC9747648-A1.
PN 18-DEC-1997.
PD 18-DEC-1997.
PF 12-JUN-1997: J02031.
PR 14-JUN-1994: JP-154527.
PA (MEIP ) MEIJI MILK PROD CO LTD.
P: Dairiki K, Kino K.
DR WPI: 98-052242/05.
PT T-cell epitope peptide portion of Japanese cypress pollen antigens
PT Chaol and Chaoc2 - used for diagnosis and treatment of spring tree
PT pollen disease
PS Claim 1: Page 32: 71pp: Japanese.
CC The present sequence represents a T-cell epitope peptide from Japanese
CC cypress pollen antigen Chaol. The present invention describes peptides
CC which correspond to the T-cell epitope sites on Japanese cypress pollen
CC antigens Chaol and Chaoc2. The peptides can be used as a reagent for the
CC diagnosis of allergy to Japanese cypress pollen, and as an antigen in
CC the treatment and prevention of spring tree pollen disease in which the
CC pollinosis involves reactivity to Japanese cypress pollen.
SQ Sequence 20 AA:

Query Match 60.4% Score 125: DB 1: Length 20:
Best Local Similarity 100.0% Pred. No. 3,41e-04:
Matches 20: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

DE 1 SSGKNEGNIYNNNEAFKVE 20
DQ 1 SSGKNEGNIYNNNEAFKVE 20

RESULT 5
ID W04344 standard: Protein: 354 AA.
AC W04344.
DT 28-NOV-1996 (first entry)
DE Chamaecyparis obtusa pollen allergen Cha o 1 (A).
KW Pollen allergen: Cha o 1: T-cell epitope, prevention: treatment,
KW pollinosis.
OS Chamaecyparis obtusa.
PN J08176192-A.
PD 09-JUL-1996.
PF 21-DEC-1994: JP-335089.
PR 21-DEC-1994: JP-335089.
PA (MEIP ) MEIJI MILK PROD CO LTD.
DR WPI: 98-368225/37.
DR N-PSDB: T38619.
PT DNA encoding chamaecyparis obtusa pollen allergen - T cell
PT epitope(s) of which are useful in development of preventative and
PT treating agent for C. obtusa pollen pollinosis
PS Claim 2: Pages 10-11: 17pp: Japanese
CC The present sequence is the C. obtusa pollen allergen Cha o 1.
CC the T-cell epitopes of which can be used in the development of a
CC preventive and treating agent for C. obtusa pollen pollinosis.
CC C. obtusa pollen (2.4 kb) was decreased with diethyl ether, and
CC dried at room temp. overnight. Cha o 1 was sepd. from it and
CC purified. RNA was extracted from C. obtusa pollen, and mRNA and
CC cDNA derived.
SQ Sequence 354 AA:

Query Match 60.4% Score 125: DB 1: Length 354:
Best Local Similarity 100.0% Pred. No. 3,41e-04:
Matches 20: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

DQ 311 SSGKNEGNIYNNNEAFKVE 352
DQ 1 SSGKNEGNIYNNNEAFKVE 20

RESULT 4
ID W42121 standard: Protein: 354 AA.
AC W42121.
DT 16-JUN-1998 (first entry)
DE Japanese cypress pollen antigen Chaoc1.
KW Japanese cypress pollen antigen: Chaoc1 epitope: Chaoc1: Chaoc2.

```

```

KW diagnosis: allergy: spring tree pollen disease: pollinosis.
OS Chamaecyparis obtusa.
PN WC9747648-A1.
PD 18-DEC-1997.
PF 12-JUN-1997: J02031.
PR 14-JUN-1994: JP-153527.
PA (MEIP ) MEIJI MILK PROD CO LTD.
P: Dairiki K, Kino K.
DR WPI: 98-052242/05.
PT T-cell epitope peptide portion of Japanese cypress pollen antigens
PT Chaol and Chaoc2 - used for diagnosis and treatment of spring tree
PT pollen disease
PS Example 1: Page 14-15: 7pp: Japanese.
CC The present sequence represents Japanese cypress pollen antigen Chaol.
CC The present invention describes peptides which correspond to the T-cell
CC epitope sites on Japanese cypress pollen antigens Chaol and Chaoc2. The
CC peptides can be used as a reagent for the diagnosis of allergy to
CC Japanese cypress pollen, and as an antigen in the treatment and
CC prevention of spring tree pollen disease in which the pollinosis
CC involves reactivity to Japanese cypress pollen.
SQ Sequence 354 AA:

Query Match 60.4% Score 125: DB 1: Length 354:
Best Local Similarity 100.0% Pred. No. 3,41e-04:
Matches 20: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

DQ 311 SSGKNEGNIYNNNEAFKVE 350
DQ 1 SSGKNEGNIYNNNEAFKVE 20

RESULT 5
ID W04345 standard: Protein: 375 AA.
AC W04345.
DT 28-NOV-1996 (first entry)
DE Chamaecyparis obtusa pollen allergen Cha o 1 cDNA (B).
KW Pollen allergen: Cha o 1: T-cell epitope, prevention: treatment,
KW pollinosis.
OS Chamaecyparis obtusa.
PN J08176192-A.
PD 09-JUL-1996.
PF 21-DEC-1994: JP-335089.
PR 21-DEC-1994: JP-335089.
PA (MEIP ) MEIJI MILK PROD CO LTD.
DR WPI: 98-368225/37.
DR N-PSDB: T38619.
PT DNA encoding chamaecyparis obtusa pollen allergen - T cell
PT epitope(s) of which are useful in development of preventative and
PT treating agent for C. obtusa pollen pollinosis
PS Claim 8: Pages 11-12: 17pp: Japanese
CC The present sequence is the C. obtusa pollen allergen Cha o 1.
CC the T-cell epitopes of which can be used in the development of a
CC preventive and treating agent for C. obtusa pollen pollinosis.
CC C. obtusa pollen (2.4 kb) was decreased with diethyl ether, and
CC dried at room temp. overnight. Cha o 1 was sepd. from it and
CC purified. RNA was extracted from C. obtusa pollen, and mRNA and
CC cDNA derived.
SQ Sequence 375 AA:

Query Match 60.4% Score 125: DB 1: Length 375:
Best Local Similarity 100.0% Pred. No. 3,41e-04:
Matches 20: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

DQ 332 SSGKNEGNIYNNNEAFKVE 352
DQ 1 SSGKNEGNIYNNNEAFKVE 20

RESULT 5
ID R45577 standard: Protein: 367 AA.
AC R45577.
DT 13-JUL-1994 (first entry)
DE Jun S I.

```


KW Japanese cedar: pollen allergen; allergy: treatment; diagnosis;
 KW T cell epitope; sensitivity; detection;
 OS Juniperus sabroides.
 FH Key Location/Qualifiers
 FI peptide 1: 21: "signal peptide"
 FI peptide /note= "mature peptide"
 FI peptide 22: 367
 FI peptide /note= "mature peptide"
 PN WO9401560-A.
 PD 20-JAN-1994.
 PE 15-JAN-1993: U00139.
 PR 10-JUL-1992: WO-935661.
 PR 01-SEP-1992: US-938390.
 PA (IMMU-) IMMULOGIC PHARM CORP.
 PI Bond JF, Garman RD, Griffith LJ, Kuo M, Pollock J.
 DK WPI: 94-035066/04.
 DR N-PS2B: 055272.
 PI Antigens derived from Japanese cedar pollen allergen Cry j 1.
 PI contain at least two T cell epitope(s), used to treat or diagnose
 PI allergy.
 PS Disclosure: Fig 16: 137pp; English.
 CC The sequence is that of Jun S I, a homologue of the Japanese
 CC cedar pollen allergen Cry j 1. Antigenic peptides derived from it
 CC can be used for the treatment and diagnosis of allergies associated
 CC with Japanese cedar pollen.
 SC Sequence 367 AA:

Query Match 52.2% Score 108: DB 1: Length 367;
 Best Local Similarity 85.0% Pred. No. 1.49e-02;
 Matches 17: Conservative 1: Mismatches 2: Indels 0: Gaps 0:

DB 332 SSGKTEETNINNSNEAFKVE 351
 IIII I I I I I I I I I
 QY 1 SSGKNEGTINYNNEAFKVE 20

RESULT 7
 ID R82522 standard: Protein: 20 AA.
 AC R82522:
 DI 15-APR-1996 (first entry)
 DE Cry j 1 Japanese Cedar pollen allergen peptide fragment (CJ1-32).
 KW Cry j 1: Japanese cedar pollen allergen; modified; drug production;
 KW allergy: Cryptomeria japonica.
 OS Cryptomeria japonica.
 PN WO9527786-A1.
 PL 19-OCT-1995.
 PF 06-APR-1995: U04249.
 PR 08-APR-1994: US-526248.
 PR 06-DEC-1994: US-552226.
 PA (IMMU-) IMMULOGIC PHARM CORP.
 PI Chen X, Evans S, Francis RM, Kuo M, Powers SF.
 PI Shaked Z.
 DK WPI: 95-36631/47.
 DR Modified Cryptomeria japonica (Cry j) 1 peptide(s) - useful for
 DR treating allergy to Japanese cedar pollen allergen or
 DR immunologically cross reactive allergens
 PS Disclosure: Figure 2: 60pp; English.
 CC Novel peptides of cry j 1 have been modified as a part of a
 CC preformulation scheme to develop an optimised drug product for
 CC therapeutic treatment of humans suffering from allergy to Japanese
 CC cedar pollen allergen or an allergen which is immunologically cross
 CC reactive with Japanese cedar pollen allergen. Such modified peptides
 CC possess certain characteristics which render them particularly
 CC suitable for drug product formulation. Peptide fragments of Cry j 1,
 CC modified and unmodified, are given in R82491-R82525. This peptide
 CC fragment corresponds to amino acids 311-330 of the allergen mature
 CC protein.
 SC Sequence 20 AA:

Query Match 44.9% Score 93: DB 1: Length 20;
 Best Local Similarity 70.0% Pred. No. 3.80e-01;
 Matches 14: Conservative 3: Mismatches 3: Indels 0: Gaps 0:

DB 1 SSGKYEGGNIYTKKEAFNVE
 IIII I I I I I I I I I
 QY 1 SSGKNEGTINYNNEAFKVE 20

RESULT 8
 ID R45573 standard: Protein: 20 AA.
 AC R45573:
 DI 13-JUL-1994 (first entry)
 DE Cry j 1 pollen allergen peptide CJ1-32
 KW Japanese cedar: detection; allergy; treatment; diagnosis;
 KW T cell epitope; sensitivity;
 OS Cryptomeria japonica.
 PN WO9401560-A.
 PD 20-JAN-1994.
 PE 15-JAN-1993: U00139.
 PR 10-JUL-1992: WO-935661.
 PR 01-SEP-1992: US-938990.
 PA (IMMU-) IMMULOGIC PHARM CORP.
 PI Bond JF, Garman RD, Griffith LJ, Kuo M, Pollock J.
 DK WPI: 94-035066/04.
 DR Antigens derived from Japanese cedar pollen allergen Cry j 1.
 DR contain at least two T cell epitope(s), used to treat or diagnose
 DR allergy.
 PS Claim 1: Fig 13: 137pp; English.
 CC The sequence is that of an isolated peptide of the Japanese cedar
 CC pollen allergen Cry j 1 (amino acids 311-330). The peptide, CJ1-32,
 CC can be used for the treatment and diagnosis of allergies associated
 CC with Japanese cedar pollen. It has enhanced therapeutic properties
 CC but reduced side effects compared to naturally occurring allergens.
 SC Sequence 20 AA:

Query Match 44.9% Score 93: DB 1: Length 20;
 Best Local Similarity 70.0% Pred. No. 3.80e-01;
 Matches 14: Conservative 3: Mismatches 3: Indels 0: Gaps 0:

DB 1 SSGKYEGGNIYTKKEAFNVE 20
 IIII I I I I I I I I I
 QY 1 SSGKNEGTINYNNEAFKVE 20

RESULT 9
 ID R83056 standard: Peptide: 20 AA.
 AC R83056:
 DI 15-MAR-1999 (first entry)
 DE Control peptide P16.
 KW Pasi epitope: HFE7A; monoclonal antibody; humanised antibody;
 KW human: apertosis: HFE7A; autoimmune disease; Hashimoto's disease;
 KW systemic lupus erythematosus; viral versus host disease;
 KW Sjogren syndrome; pernicious anaemia; Addison's disease;
 KW scleroderma; Goodpasture syndrome; Crohn's disease; sterility;
 KW rheumatoid arthritis; autoimmune haemolytic anaemia;
 KW myasthenia gravis; multiple sclerosis; Basedow's disease;
 KW thrombopenia purpura; insulin-dependent diabetes; allergy;
 KW atopy; arteriosclerosis; myocarditis; cardiomyopathy; AIDS;
 KW glomerular nephritis; hypoplastic anaemia; hepatitis; AIDS;
 KW transplant rejection; therapy.
 OS Synthetic.
 PN AU9859701-A.
 PD 08-OCT-1998.
 PF 30-MAR-1998: 059701.
 PR 08-OCT-1997: JP-276064.
 PR 01-APR-1997: JP-082953.
 PR 25-JUN-1997: JP-169088.
 PA (SANY) SANKYO CO LTD.
 PI Akio S, Hideyuki H, Hiroko Y, Jun O, Kimihisa I,
 PI Masahiko O, Nobufusa S, Shin Y, Tohru T.
 DK WPI: 98-54340/47.

PT New antibodies and proteins bind conserved epitope of Ras antigen -
 PT used to evaluate drugs in animal models and to treat Ras-associated
 PT diseases e.g. autoimmune disease, allergy, atopy, arteriosclerosis,
 PT myocarditis, hepatitis and AIDS.
 PS Reference Example 6; Page 87; 292pp; English.

CC Synthetic peptides PL-P15 (see W83043-57) are partial sequences of
 CC amino acids 1-157 of the extracellular domain of human Fas, with
 CC between 9 and 11 amino acid residues overlapping one another. P16
 CC (see W83058) is a negative control having no homology with human
 CC Fas. P1-P15 were used in an ELISA, which demonstrated that novel
 CC murine anti-human Fas monoclonal antibody HFE7A specifically binds
 CC an amino acid sequence contained in P11. The epitope (see W83030)
 CC was subsequently identified. The invention provides humanised
 CC HFE7A antibodies (see W83041-57) produced by CDR grafting. These
 CC antibodies are capable of inducing apoptosis in abnormal cells
 CC expressing Fas, and of inhibiting Fas-induced apoptosis in normal
 CC cells. They are used to evaluate, in animal models, treatments of
 CC diseases that involve Fas/Fas ligand interactions, and also to treat
 CC such diseases.
 CC Sequence 20 AA:

Query Match 44.9% Score 93; DB 1; Length 20;

Best Local Similarity 70.0%; Pred. No. 3.80e-01;

Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 1 SSGKYECCGNIYTKKEAFNVE 20

QY 1 SSGKNEGTNIYNNNEAFKVE 20

RESULT 12

ID W80355 standard; peptide; 21 AA.

AC W80355

DE 11-JAN-1999 (first entry)

CC Sugi allergen protein Cry2 derived epitope for T cells.

KW T cell epitope; sugi allergen proteins Cry1; Cry2; treatment;

KW sugi-pollinosis; allergic reaction; pollen;

OS Synthetic.

PN J10259188-A.

PD 22-SEP-1998.

PF 22-DEC-1997; 353448.

PR 24-DEC-1996; JP-34341.

PA (HAYE) HAYASHIBARA SEIBUITSU KASAKU.

FA (SANY) SANKYO CO LTD.

DR WPI: 98-577037/49.

PT A linked T cell epitope peptide - used for the treatment of

PT Sugi-pollinosis.

PS Claim 8; pages 19-20; 21pp; Japanese.

CC W80354-58 represent epitopes for T cells, derived from the sugi allergen

CC proteins Cry1 (W80339-44, W80350-53 and W80355-58) and Cry2 (W80345-49

CC and W80344-56). The peptides are useful for the treatment of

CC sugi-pollinosis, an allergic reaction of the body to pollen.

CC Sequence 21 AA:

Query Match 44.9% Score 93; DB 1; Length 20;

Best Local Similarity 70.0%; Pred. No. 3.80e-01;

Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 2 SSGKYECCGNIYTKKEAFNVE 21

QY 1 SSGKNEGTNIYNNNEAFKVE 20

RESULT 13

ID R45598 standard; Protein; 26 AA.

AC R45598

DE 13-JUL-1994 (first entry)

CC Cry J I pollen allergen peptide CJI-44.

KW Japanese cedar; detection; allergy; treatment; diagnosis;

KW T cell epitope; sensitivity.

OS Cryptomeria japonica.

PN W89401560-A.

PD 20-JAN-1994.

PF 15-JAN-1993; U00139.

PR 10-JUL-1992; WO-0035661.

FA (IMMUG) IMMUGLOGIC PHARM CORP.

DR Bond JF, Garman RD, Griffith IJ, Kuo M, Pollock J.

DR WPI: 94-035055/04.

PT Antigens derived from Japanese cedar pollen allergen Cry J I

PT contain at least two T cell epitope(s), used to treat or diagnose

PT allergy.

PS Claim 76; Fig 18; 137pp; English.

CC The sequence is that of an isolated peptide of the Japanese cedar

CC pollen allergen Cry J I. The peptide, CJI-44.1, can be used for

CC the treatment and diagnosis of allergies associated with Japanese

CC cedar pollen. It has enhanced therapeutic properties but reduced

CC side effects compared to naturally occurring allergens.

CC Sequence 26 AA:

Query Match 44.9% Score 93; DB 1; Length 26;

Best Local Similarity 70.0%; Pred. No. 3.80e-01;

Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 7 SSGKYECCGNIYTKKEAFNVE 26

QY 1 SSGKNEGTNIYNNNEAFKVE 20

RESULT 12

ID R92180 standard; Protein; 28 AA.

AC R92180

DE 16-APR-1996 (first entry)

CC Cry J I Japanese Cedar pollen allergen modified peptide (CJI-44.8)

KW Cry J I; Japanese cedar pollen allergen; modified; drug production;

KW allergy; Cryptomeria japonica.

OS Synthetic.

PN W89527786-A.

PD 19-OCT-1995.

PF 06-APR-1995; U04249.

PR 08-APR-1994; US-226248.

PR 06-DEC-1994; US-350225.

PA (IMMUG) IMMUGLOGIC PHARM CORP.

PI Chen X, Evans S, Franzen HM, Kuo M, Powers SP.

PI Shaked Z.

PI WPI: 95-366392/47.

DR Modified Cryptomeria japonica (Y J) I peptides) - useful for

DR treating allergy to Japanese cedar pollen allergen or

PT immunologically cross reactive allergens

PS Claim 1; Figure 3; 60pp; Eng.

CC Novel peptides of cry J I have been modified as a part of a

CC preformulation scheme to develop an optimized drug product for

CC therapeutic treatment of humans suffering from allergy to Japanese

CC cedar pollen allergen or an allergen which is immunologically cross

CC reactive with Japanese cedar pollen allergen. Such modified peptides

CC possess certain characteristics which render them particularly

CC suitable for drug product formulation. In the treatment of

CC modified and unmodified are then used to treat the allergic

CC is one of three modified per day (see W89527786)

CC Sequence 28 AA:

Query Match 44.9% Score 93; DB 1; Length 26;

Best Local Similarity 70.0%; Pred. No. 3.80e-01;

Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 9 SSGKYECCGNIYTKKEAFNVE 26

QY 1 SSGKNEGTNIYNNNEAFKVE 20

RESULT 13

ID R45598 standard; Protein; 30 AA.

AC R45598

DE 13-JUL-1994 (first entry)

CC Cry J I pollen allergen peptide CJI-44.

KW Japanese cedar; detection; allergy; treatment; diagnosis;

KW T cell epitope; sensitivity.

OS Cryptomeria japonica.

PN W09401560-A.

PD 20-JAN-1994.

PF 15-JAN-1993; U00139.

```

PR 10-JUL-1992; WO-905661.
PP 01-SEP-1992; US-938990.
PA (IMMUNO) IMMUNOLOGIC PHARM CORP
PI Bond JP, Garman RD, Griffith LJ, Kuo M, Pollock J
CP WPI: 94-03556/04.
PT Antigen derived from Japanese cedar pollen allergen Cry j 1
PI contain at least two T-cell epitope(s), used to treat or diagnose
PI allergy.
PS Claim 76; Fig 18; 137pp; English.
CC The sequence is that of an isolated peptide of the Japanese cedar
CC pollen allergen Cry j 1. The peptide, C31-44, can be used for
CC the treatment and diagnosis of allergies associated with Japanese
CC cedar pollen. It has enhanced therapeutic properties but reduced
CC side effects compared to naturally occurring allergens.
CC Sequence 353 AA:
DB 11 SSKYEGGNIYTKKAFNVE 30
QY 1 SSKKNEGNIYNNNAKVE 20

RESULT 14
IL R75388 standard; protein: 353 AA.
AC R75388:1995 (first entry)
DE Japanese cedar pollen allergen Cry j 1.
KW Japanese cedar pollen allergen Cry j 1; T-cell epitope; peptides.
KW prevention; treatment; cryptomeria pollinosis.
CS Cryptomeria japonica.
PE Key Location/Qualifiers
PT peptide 61-175
P1 /note="T-cell epitope peptide"
P2 peptide 91-105
P3 peptide 106-120
P4 peptide 146-160
P5 peptide 211-225
P6 peptide 326-340
P7 peptide 355-445
P8 C0711d295-A.
PD 09-MAY-1995
PF 20-OCT-1993; 254426.
PP 20-OCT-1993; JP-262127.
PA (MEIJ) MEIJT MALK 8900 Co. LTD.
DB WPI: 95-201834/27.
PI New cryptomeria pollen allergen T-cell epitope peptide - used for
PI prevention, treatment and investigation of Japanese cedar pollinosis
PS Disclosed: Figs 1-2, 8pp; Japanese.
CC R75388 is the Japanese cedar pollen allergen Cry j 1, from which the
CC T-cell epitope peptides R89289-R89295 were derived. The peptides
CC can be used for the prevention and treatment of cryptomeria
CC pollinosis, and also for the investigation of pollinosis.
CC Sequence 353 AA:
Query Match 44.9% Score 93; DB 1; Length 353;
Best Local Similarity 70.0%; Pred. No. 3,80e-01;
Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

DB 311 SSKYEGGNIYTKKAFNVE 330
QY 1 SSKKNEGNIYNNNAKVE 20

Query Match 44.9% Score 93; DB 1; Length 353;
Best Local Similarity 70.0%; Pred. No. 3,80e-01;
Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
Search completed: Mon Jun 19 16:03:42 2000
CPU time : 7 secs.

```

```

ID R81587 standard; protein: 353 AA.
AC R81587:1995 (first entry)
DE Cedar pollen allergen B.
KW Cedar pollen allergen; immunoglobulin E; IgE; T-cell epitope
KW antibody; pollinosis; therapy; immunotherapy.
QS Cryptomeria japonica.
PN EP-700929-A2.
PD 13-MAR-1996.
PF 08-SEP-1995; 306295.
PR 10-SEP-1994; JP-242137.
PR 14-JUL-1995; JP-200221.
PR 14-JUL-1995; JP-200324.
PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
PI Hino K, Saito S, Taniguchi Y.
DR WPI: 96-140975/15.
PT New peptide(s) derived from cedar pollen allergens - activate
PT allergen-specific T-cells, but not allergen-specific IgE antibody.
PI used for treating cedar pollinosis.
PS Claim 5; Page 31-32; 36pp; English.
CC Synthetic peptides based on portions of cedar pollen allergens A
CC (R81586) and B (R81587) were tested for their ability to activate
CC cedar allergen-specific T-cells, but not allergen-specific IgE
CC antibodies. 6 peptides (R81588-R81595) were identified as T-cell
CC epitopes. These peptides, plus subsequences (R81573-79) essential
CC for T-cell recognition, and homologous peptides (R81588-96) can
CC be used as immunotherapeutic agents to treat or prevent cedar
CC pollinosis, avoiding side-effects such as anaphylaxis.
CC Sequence 353 AA:
Query Match 44.9% Score 93; DB 1; Length 353;
Best Local Similarity 70.0%; Pred. No. 3,80e-01;
Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

DB 311 SSKYEGGNIYTKKAFNVE 330
QY 1 SSKKNEGNIYNNNAKVE 20

Search completed: Mon Jun 19 16:03:42 2000
CPU time : 7 secs.

```


| | | |
|--------|--|---|
| CC | TOPOLOGY: | LINEAR |
| CC | M-LECULE TYPE: | PROTEIN |
| SQ | SEQUENCE | 317 AA: 37355 MW: 525334 CN: |
| | Query Match | 34.8%; Score 72; DB 2; Length 317; |
| | Best Local Similarity | 40.0%; Pred. No. 1.69e-01; |
| | Matches | 10; Conservative 7; Mismatches 7; Indels 1; Gaps 1; |
| DB | 57 | STHAYAPP-PYKIMDRFFKRVDVI 80 |
| QY | 7 | GTYNNNEAFKVERFIKRVSNV 31 |
| RESULT | 2 | |
| ID | US-08-459-263-2 | STANDARD: PRT: 317 AA. |
| XX | XXXXXX | |
| DE | Sequence 2, Application US/08459263 | |
| CC | Patient No. 5925749 | |
| CC | GENERAL INFORMATION: | |
| CC | APPLICANT: Mathur, E., et al. | |
| CC | TITLE OF INVENTION: Carboxymethyl Cellulase from Thermotoga Maritima | |
| CC | NUMBER OF SEQUENCES: 4 | |
| CC | CORRESPONDENCE ADDRESS: | |
| CC | ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, | |
| CC | ADDRESS: CECCHI, STEWART & OLSTEIN | |
| CC | STREET: 6 BECKER FARM ROAD | |
| CC | CITY: ROSELAND | |
| CC | STATE: NEW JERSEY | |
| CC | COUNTRY: USA | |
| CC | ZIP: 07068 | |
| CC | COMPUTER READABLE FORM: | |
| CC | MEDIUM TYPE: 3.5 INCH DISKETTE | |
| CC | COMPUTER: IBM PS/2 | |
| CC | OPERATING SYSTEM: MS-DOS | |
| CC | SOFTWARE: WORD PERFECT 5.1 | |
| CC | CURRENT APPLICATION DATA: | |
| CC | APPLICATION NUMBER: US/08459263 | |
| CC | FILING DATE: | |
| CC | CLASSIFICATION: | |
| CC | PRIOR APPLICATION DATA: | |
| CC | APPLICATION NUMBER: US/08459263 | |
| CC | FILING DATE: August 23, 1995 | |
| CC | ATTORNEY/AGENT INFORMATION: | |
| CC | NAME: FERRARO, GREGORY D. | |
| CC | REGISTRATION NUMBER: 36114 | |
| CC | REFERENCE/DOCKET NUMB: 331400-20 | |
| CC | TELECOMMUNICATION INFORMATION: | |
| CC | TELEPHONE: 201-994-1700 | |
| CC | TELEFAX: 201-994-1744 | |
| CC | INFORMATION FOR SEQ ID NO: 2: | |
| CC | SEQUENCE CHARACTERISTICS: | |
| CC | LENGTH: 317 AMINO ACIDS | |
| CC | TYPE: AMINO ACID | |
| CC | STRANDEDNESS: | |
| CC | TOPOLOGY: LINEAR | |
| CC | M-LECULE TYPE: PROTEIN | |
| SQ | SEQUENCE | 317 AA: 37355 MW: 525334 CN: |
| | Query Match | 34.8%; Score 72; DB 3; Length 317; |
| | Best Local Similarity | 40.0%; Pred. No. 1.69e-01; |
| | Matches | 10; Conservative 7; Mismatches 7; Indels 1; Gaps 1; |
| DB | 57 | STHAYAPP-PYKIMDRFFKRVDVI 80 |
| QY | 7 | GTYNNNEAFKVERFIKRVSNV 31 |

CC APPLICANT: RASMUSSEN, Beth A
CC APPLICANT: TALLY, Francis P
CC APPLICANT: GLUZMAN, Yakov
CC TITLE OF INVENTION: DNA SEQUENCES AND AMINO ACID SEQUENCES
CC TITLE OF INVENTION: OF CLASS B BETA-LACTAMASE ENZYMES FROM BACTEROIDES
CC TITLE OF INVENTION: FRAGILIS
CC NUMBER OF SEQUENCES: 9
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Darby & Darby PC
CC STREET: 805 Third Avenue
CC CITY: New York
CC STATE: NY
CC COUNTRY: US
CC ZIP: 10022
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/459,263
CC FILING DATE: 02-JUN-1995
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Robinson, Joseph R
CC REGISTRATION NUMBER: 33,448
CC REFERENCE/DOCKET NUMBER: 0646/18026-US2
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 212-527-7700
CC TELEFAX: 212-753-6237
CC TELEX: 236687
CC INFORMATION FOR SEQ ID NO: 4:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 255 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC ORIGINAL SOURCE:
CC ORGANISM: Bacillus cereus
CC IMMEDIATE SOURCE:
CC CLONE: P-LACTAMASE
SQ SEQUENCE 255 AA: 27847 MW: 349604 CN:

Query Match 34.38; Score 71; DB 1; Length 255;
Best Local Similarity 57.18; Pct. No. 2.06e-01;
Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

DE 98 MVEKKFKRVTDVI 111
QY 18 KVERRFKRVSNV 31

RESULT 5
ID US-08-459-264-4 STANDARD: PRT: 255 AA.
XX
AC xxxxxx
XX
DT
DE Sequence 4, Application US/08459264
XX Sequence 4, Application US/08459264
CC Patent No. 5705340
CC GENERAL INFORMATION:
CC APPLICANT: RASMUSSEN, Beth A
CC APPLICANT: TALLY, Francis P
CC APPLICANT: GLUZMAN, Yakov
CC TITLE OF INVENTION: DNA SEQUENCES AND AMINO ACID SEQUENCES
CC TITLE OF INVENTION: OF CLASS B BETA-LACTAMASE ENZYMES FROM BACTEROIDES
CC TITLE OF INVENTION: FRAGILIS
CC NUMBER OF SEQUENCES: 9
CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Darby & Darby PC
CC STREET: 805 Third Avenue
CC CITY: New York
CC STATE: NY
CC COUNTRY: US
CC ZIP: 10022
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/459,264
CC FILING DATE: 02-JUN-1995
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Robinson, Joseph R
CC REGISTRATION NUMBER: 33,448
CC REFERENCE/DOCKET NUMBER: 0546/18026-US3
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 212-527-7700
CC TELEFAX: 212-753-6237
CC TELEX: 236687
CC INFORMATION FOR SEQ ID NO: 4:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 255 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC ORIGINAL SOURCE:
CC ORGANISM: Bacillus cereus
CC IMMEDIATE SOURCE:
CC CLONE: B-LACTAMASE
SQ SEQUENCE 255 AA: 27847 MW: 349604 CN:

Query Match 34.38; Score 71; DB 1; Length 255;
Best Local Similarity 57.18; Pct. No. 2.06e-01;
Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

DE 98 MVEKKFKRVTDVI 111
QY 18 KVERRFKRVSNV 31

RESULT 6
ID 55:4582-41 STANDARD: PRT: 255 AA.
XX
AC xxxxxx
XX
DT
DE Patent No. 5514582
XX Patent No. 5514582
CC APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.
CC TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID
CC IMMUNOGLOBULINS
CC NUMBER OF SEQUENCES: 43
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/185,670
CC FILING DATE: 21-JAN-1994
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 986,931
CC FILING DATE: 08-DEC-1992
CC APPLICATION NUMBER: 603,122
CC FILING DATE: 16-DEC-1991
CC APPLICATION NUMBER: 440,625
CC FILING DATE: 22-NOV-1989
CC APPLICATION NUMBER: 315,015
CC FILING DATE: 23-FEB-1989
CC SEQ ID NO: 41:
CC LENGTH: 86

WATERMAN

(TM)

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (C) 1993-1998 University of Edinburgh, U.K.
Distribution Rights by Oxford Molecular Ltd

MPstrch_app protein - protein database search, using Smith-Waterman algorithm
Run On: Mon Jun 19 16:23:02 2000 MasPar time 16.08 Seconds
194.302 Million cell updates/sec
Tabular output not generated.

Title: >US-09-142-524A-5
Description: (1-31) from US09-142524A.pep
Perfect Score: 207
Sequence: 1 SSKNEGNIYNNNEAFKVERRF<KRVSNVI 31

Scoring table: FAM 150
Gap 11

Searched: 721208 seqs, 100765575 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-pending
1:PCI 2:J6 3:G0 4:U7 5:G0 6:U8 7:U82 8:U83 9:U84
10:U84R 11:U85 12:U86 13:U87 14:U88 15:U89 16:U90 17:U91
18:U92 19:U93 20:U94 21:U95 22:NEAP 23:NEWC60 24:NEW08
25:NEW09

Statistics: Mean 25.118; Variance 80.992; scale 0.310

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description | Pred. No. |
|------------|-------|-------------|--------|------------|------------------------|-----------|
| 1 | 207 | 100.0 | 31 | US-09-142- | Sequence 5, Applicatio | 5.98e-01 |
| 2 | 108 | 92.2 | 346 | US-62-109- | Sequence 1, Applicatio | 1.77e-02 |
| 3 | 108 | 92.2 | 367 | US-08-226- | Sequence 95, Applicati | 1.77e-02 |
| 4 | 108 | 92.2 | 367 | US-08-467- | Sequence 95, Applicati | 1.77e-02 |
| 5 | 108 | 92.2 | 367 | US-08-467- | Sequence 95, Applicati | 1.77e-02 |
| 6 | 108 | 92.2 | 367 | US-08-350- | Sequence 95, Applicati | 1.77e-02 |
| 7 | 108 | 92.2 | 367 | US-08-468- | Sequence 95, Applicati | 1.77e-02 |
| 8 | 108 | 92.2 | 367 | US-08-468- | Sequence 95, Applicati | 1.77e-02 |
| 9 | 97 | 45.9 | 30 | US-08-467- | Sequence 261, Applicat | 2.60e-01 |
| 10 | 97 | 45.9 | 30 | US-08-350- | Sequence 261, Applicat | 2.60e-01 |
| 11 | 97 | 45.9 | 30 | US-08-467- | Sequence 261, Applicat | 2.60e-01 |
| 12 | 97 | 45.9 | 30 | US-08-468- | Sequence 261, Applicat | 2.60e-01 |
| 13 | 97 | 45.9 | 30 | US-08-468- | Sequence 261, Applicat | 2.60e-01 |
| 14 | 93 | 44.9 | 20 | US-08-468- | Sequence 57, Applicati | 6.84e-01 |
| 15 | 93 | 44.9 | 20 | US-07-938- | Sequence 57, Applicati | 6.84e-01 |
| 16 | 93 | 44.9 | 20 | US-08-350- | Sequence 57, Applicati | 6.84e-01 |
| 17 | 93 | 44.9 | 26 | US-08-226- | Sequence 91, Applicati | 6.84e-01 |
| 18 | 93 | 44.9 | 26 | US-08-350- | Sequence 91, Applicati | 6.84e-01 |
| 19 | 93 | 44.9 | 25 | US-08-467- | Sequence 91, Applicati | 6.84e-01 |
| 20 | 93 | 44.9 | 28 | US-08-468- | Sequence 132, Applicat | 6.84e-01 |

| | | | | | | | |
|----|----|------|-----|----|------------|------------------------|----------|
| 21 | 93 | 44.9 | 28 | 10 | US-08-467- | Sequence 132, Applicat | 6.84e-01 |
| 22 | 93 | 44.9 | 28 | 10 | US-08-467- | Sequence 132, Applicat | 6.84e-01 |
| 23 | 93 | 44.9 | 28 | 10 | US-08-467- | Sequence 131, Applicat | 6.84e-01 |
| 24 | 93 | 44.9 | 29 | 10 | US-08-467- | Sequence 131, Applicat | 6.84e-01 |
| 25 | 93 | 44.9 | 29 | 10 | US-08-467- | Sequence 131, Applicat | 6.84e-01 |
| 26 | 93 | 44.9 | 30 | 10 | US-08-467- | Sequence 260, Applicat | 6.84e-01 |
| 27 | 93 | 44.9 | 30 | 10 | US-08-467- | Sequence 260, Applicat | 6.84e-01 |
| 28 | 93 | 44.9 | 30 | 10 | US-08-467- | Sequence 90, Applicati | 6.84e-01 |
| 29 | 93 | 44.9 | 30 | 10 | US-08-467- | Sequence 90, Applicati | 6.84e-01 |
| 30 | 93 | 44.9 | 30 | 10 | US-08-467- | Sequence 90, Applicati | 6.84e-01 |
| 31 | 93 | 44.9 | 30 | 10 | US-08-468- | Sequence 90, Applicati | 6.84e-01 |
| 32 | 93 | 44.9 | 30 | 10 | US-08-468- | Sequence 90, Applicati | 6.84e-01 |
| 33 | 93 | 44.9 | 40 | 8 | US-08-350- | Sequence 70, Applicati | 6.84e-01 |
| 34 | 93 | 44.9 | 40 | 8 | US-08-350- | Sequence 70, Applicati | 6.84e-01 |
| 35 | 93 | 44.9 | 40 | 8 | US-08-350- | Sequence 70, Applicati | 6.84e-01 |
| 36 | 93 | 44.9 | 40 | 8 | US-08-350- | Sequence 70, Applicati | 6.84e-01 |
| 37 | 93 | 44.9 | 40 | 8 | US-08-350- | Sequence 70, Applicati | 6.84e-01 |
| 38 | 93 | 44.9 | 40 | 8 | US-08-350- | Sequence 70, Applicati | 6.84e-01 |
| 39 | 93 | 44.9 | 63 | 10 | US-08-467- | Sequence 55, Applicati | 6.84e-01 |
| 40 | 93 | 44.9 | 63 | 10 | US-08-467- | Sequence 55, Applicati | 6.84e-01 |
| 41 | 93 | 44.9 | 63 | 10 | US-08-467- | Sequence 65, Applicati | 6.84e-01 |
| 42 | 93 | 44.9 | 63 | 10 | US-08-467- | Sequence 65, Applicati | 6.84e-01 |
| 43 | 93 | 44.9 | 353 | 11 | US-08-526- | Sequence 15, Applicati | 6.84e-01 |
| 44 | 93 | 44.9 | 374 | 10 | US-08-468- | Sequence 2, Applicatio | 6.84e-01 |
| 45 | 93 | 44.9 | 374 | 4 | US-07-729- | Sequence 2, Applicatio | 6.84e-01 |

ALIGNMENTS

RESULT 1
ID US-09-142-524-5 STANDARD: PRI 1 AA
XX xxxxxx

DT
XX

DE Sequence 5, Application: US/09142524

XX Sequence 5, Application: US/09142524

CC GENERAL INFORMATION:

CC APPLICANT: Kume, Akino

CC APPLICANT: Kume, Akino

CC APPLICANT: Kume, Akino

CC APPLICANT: Kume, Akino

CC APPLICANT: Kume, Akino

CC APPLICANT: Kume, Akino

CC APPLICANT: Kume, Akino

CC APPLICANT: Kume, Akino

CC APPLICANT: Kume, Akino

CC APPLICANT: Kume, Akino

CC APPLICANT: Kume, Akino

CC APPLICANT: Kume, Akino

CC APPLICANT: Kume, Akino

CC APPLICANT: Kume, Akino

CC APPLICANT: Kume, Akino

CC APPLICANT: Kume, Akino

CC APPLICANT: Kume, Akino

CC APPLICANT: Kume, Akino

CC APPLICANT: Kume, Akino

CC APPLICANT: Kume, Akino

CC APPLICANT: Kume, Akino

Query Match 100.08; Score 207; DB 17; Length 31;
Best Local Similarity 100.08; Fed. No. 5.98e-14;
Matches 31; Conservative ; Mismatches 0; Indels 0; Gaps 0;

Db 1 SSKNEGNIYNNNEAFKVERRF<KRVSNVI 31

Oy 1 SSKNEGNIYNNNEAFKVERRF<KRVSNVI 31

CC FILING DATE: December 6, 1994
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Jane E. Remillard
 CC REGISTRATION NUMBER: 38,872
 CC REFERENCE/DOCKET NUMBER: 025.6 US02 (IMI-128CPD2)
 CC TELEPHONE: (617) 227-7400
 CC TELEFAX: (617) 227-5941
 CC INFORMATION FOR SEQ ID NO: 95:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 367 amino acids
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC SEQUENCE 367 AA: 39846 MW: 695078 CN:

Query Match 52.2% Score 108 PB 10 Length 367
 Best Local Similarity 85.0% Pred. No. 1.73e-02
 Matches 17: Conservative 1: Mismatches 2: Indels 0: Gaps 0:

DB 332 SSKKTEETINYSNEAFKVE 351
 IIII IIII IIII
 OY 1 SSKKNEGNIYNNNEAFKVE 20

RESULT 5 STANDARD: PRI: 367 AA.
 ID US-08-467-006-95
 XX
 AC xxxxxx
 XX
 DE

Sequence 95, Application US/08467006

Sequence 95, Application: US/08467006

CC GENERAL INFORMATION:
 CC APPLICANT: Griffith, Irwin J.
 CC APPLICANT: Pollock, Joanne
 CC APPLICANT: Bond, Julian P.
 CC APPLICANT: Garman, Richard D.
 CC APPLICANT: Kuo, Mei-Chang
 CC APPLICANT: Yeung, Siu-mei H.
 CC APPLICANT: Brauer, Andrew
 CC APPLICANT: Exley, Mark A.
 CC APPLICANT: Powers, Steven P.
 CC TITLE OF INVENTION: Allergic Proteins And Peptides From
 CC TITLE OF INVENTION: Japanese Cedar Pollen
 CC NUMBER OF SEQUENCES: 26
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
 CC STREET: 610 Lincoln St
 CC CITY: Waltham
 CC STATE: MA
 CC COUNTRY: USA
 CC ZIP: 02154

CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patent Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/457,006
 CC FILING DATE: June 6, 1995
 CC CLASSIFICATION: 424

CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/350,225
 CC FILING DATE: December 6, 1994
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Jane E. Remillard
 CC REGISTRATION NUMBER: 38,872
 CC REFERENCE/DOCKET NUMBER: 025.6 US06 (IMI-128CPD6)
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (617) 227-7400

CC TELEFAX: (617) 227-5941
 CC INFORMATION FOR SEQ ID NO: 95:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 367 amino acids
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC SEQUENCE 367 AA: 39846 MW: 695078 CN:

Query Match 52.2% Score 108 PB 10 Length 367
 Best Local Similarity 85.0% Pred. No. 1.73e-02
 Matches 17: Conservative 1: Mismatches 2: Indels 0: Gaps 0:

DB 332 SSKKTEETINYSNEAFKVE 351
 IIII IIII IIII
 OY 1 SSKKNEGNIYNNNEAFKVE 20

RESULT 6 STANDARD: PRI: 367 AA.
 ID US-08-467-697-95
 XX
 AC xxxxxx
 XX
 DE

Sequence 95, Application US/08467697

Sequence 95, Application US/08467697

CC GENERAL INFORMATION:
 CC APPLICANT: Griffith, Irwin J.
 CC APPLICANT: Pollock, Joanne
 CC APPLICANT: Bond, Julian P.
 CC APPLICANT: Garman, Richard D.
 CC APPLICANT: Kuo, Mei-Chang
 CC APPLICANT: Yeung, Siu-mei H.
 CC APPLICANT: Brauer, Andrew
 CC APPLICANT: Exley, Mark A.
 CC APPLICANT: Powers, Steven P.
 CC TITLE OF INVENTION: Allergic Proteins And Peptides From
 CC TITLE OF INVENTION: Japanese Cedar Pollen
 CC NUMBER OF SEQUENCES: 26
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
 CC STREET: 610 Lincoln St
 CC CITY: Waltham
 CC STATE: MA
 CC COUNTRY: USA
 CC ZIP: 02154

CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patent Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/467,597
 CC FILING DATE: June 6, 1995
 CC CLASSIFICATION: 424

CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/350,225
 CC FILING DATE: December 6, 1994
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Jane E. Remillard
 CC REGISTRATION NUMBER: 38,872
 CC REFERENCE/DOCKET NUMBER: 025.6 US04 (IMI-028CPD4)
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (617) 227-7400
 CC TELEFAX: (617) 227-5941
 CC INFORMATION FOR SEQ ID NO: 95:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 367 amino acids
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein

SQ SEQUENCE 367 AA: 39846 MW: 695078 CN:

```
Query Match      52.28; Score 108; DB 10; Length 367;
Best Loca: Similarity 85.08; Pred. No. 1,738-32;
Matches 17; Conservative 2; Mismatches 2; Indels
```

00 332 SSCKTTEETNIYNSNEAFKVE 351
| | | | | : | | . | |
01 1 SSCKNESTNIYNNNEAFKVE 20

RESULT ?
TC US-06-350-225-95
STANDARD: 367 AA
PRT: 367 AA

AC

Sequence: 95. Application: 35/08350225

Sequence 95, Application US/08355225

GENERAL INFORMATION:
 APPLICANT: Griffith, Irwin J.;
 APPLICANT: Pollock, Joanne;
 APPLICANT: Reed, Julian F.;
 APPLICANT: Garman, Richard D;
 APPLICANT: Kuo, Mei-Chang;
 APPLICANT: Young, Stuart H.;
 APPLICANT: Brauer, Andrea;
 APPLICANT: Ex-cy, Mark A.;
 APPLICANT: Powers, Steven P.
 TITLE OF INVENTION: Allergenic Proteins And Peptides From
 TITLE OF INVENTION: Japanese Cedar Pollen
 NUMBER OF SEQUENCES: 26;
 CORRESPONDENCE ADDRESS:

| Query Match | Score 108: | DB 6: | Length: 367; |
|-------------|------------|-------|--------------|
| 52.28: | | | |

Best Local Similarity 85.0%; Pred. No. 1.73e-02;
Matches 17; Conservative 1; Mismatches 2

DB 332 SSGKTEETNIYNSNEAFKVE 351
 QV 333 SSGKNEGTNIYNNNEAFKVE 29

| RESULT | 8 | STANDARD | PRT |
|--------|------------------|----------|-----|
| ID | US-08-468-94C-95 | 367 AA | |

AC
XXXXXX

Sequence 95, Application: US/2016/0194000

Sequence 95, Application US/084655940
GENERAL INFORMATION:
APPLICANT: Griffith, Edwin S.
APPLICANT: Pollock, Joanne
APPLICANT: Bond, Julian F.
APPLICANT: Bond, Richard D.
APPLICANT: Kuo, Wei-Chan
APPLICANT: Yeung, Siu-Mei H.
APPLICANT: Brauer, Andrew
APPLICANT: Exley, Mark A.
APPLICANT: Powers, Steven P
TITLE OF INVENTION: Allergenic Proteins And Peptides from
TITLE OF INVENTION: Japanese Cedar Pollen
NUMBER OF SEQUENCES: 261
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.

```
Query Match      52.2%; Score 108; DB 10; Length 367;
Best Local Similarity 85.0%; Pred. No. 1,73e-02;
```

Matches 17: Conservative 1: Mismatches 2: Indels 0: Gaps 0:

DE 332 SSGKISEYNIYNNNEAFKVE 351
QY 1 SSGKNEGNIYNNNEAFKVE 20

RESULT 9
ID US-09-467-023-261 STANDARD: PRT: 30 AA

XX xxxxxx

Sequence 261: Application US/08467023

Sequence 261: Application US/08467023

GENERAL INFORMATION:
APPLICANT: Griffith, Irwin J.
APPLICANT: Pollock, Joanne
APPLICANT: Bond, Julian F.
APPLICANT: Garman, Richard D.
APPLICANT: Kuo, Mei-Chang
APPLICANT: Yeung, Siu-mei H.
APPLICANT: Brauer, Andrew
APPLICANT: Exley, Mark A.
APPLICANT: Powers, Steven P.
TITLE OF INVENTION: Allergenic Proteins And Peptides From
TITLE OF INVENTION: Japanese Cedar Pollen
NUMBER OF SEQUENCES: 261
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
STREET: 610 Lincoln St
CITY: Waltham
STATE: MA
COUNTRY: USA
ZIP: 02154

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08467023
FILING DATE: June 6, 1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/350,225
FILING DATE: December 6, 1994

ATTORNEY/AGENT INFORMATION:
NAME: John E. Remillard
REGISTRATION NUMBER: 38,872
REFERENCE/DOCKET NUMBER: 025.6 USD2 (IM1-025CPD2)
TELEPHONE: (617) 227-7430
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 261:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal

SQ SEQUENCE 30 AA: 3375 MW: 4852 CN:

Query Match 46.9% Score 97: DB 10: Length 30:

Best Local Similarity 56.7% Pred. No. 2,60e-01:

Matches 14: Conservative 4: Mismatches 3: Indels 0: Gaps 0:

DB 9 SSGKEGGNIYTKKEAFNVEK 29

QY 1 SSGKNEGNIYNNNEAFKVER 21

RESULT 11

RESULT 10

ID US-08-350-225-261 STANDARD: PRT: 30 AA

XX xxxxxx

XX

XX

Sequence 261: Application US/08350225

GENERAL INFORMATION:
APPLICANT: Griffith, Irwin J.
APPLICANT: Pollock, Joanne
APPLICANT: Bond, Julian F.
APPLICANT: Garman, Richard D.
APPLICANT: Kuo, Mei-Chang
APPLICANT: Yeung, Siu-mei H.
APPLICANT: Brauer, Andrew
APPLICANT: Exley, Mark A.
APPLICANT: Powers, Steven P.
TITLE OF INVENTION: Allergenic Proteins And Peptides From
TITLE OF INVENTION: Japanese Cedar Pollen
NUMBER OF SEQUENCES: 261
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
STREET: 610 Lincoln St
CITY: Waltham
STATE: MA
COUNTRY: USA
ZIP: 02154

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08350225
FILING DATE: December 6, 1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/226,248
FILING DATE: April 8, 1994
APPLICATION NUMBER: 07/938,990
FILING DATE: September 1, 1992
APPLICATION NUMBER: 07/0693/00139
FILING DATE: January 15, 1993

ATTORNEY/AGENT INFORMATION:
NAME: Carlene A. Vassilone
REGISTRATION NUMBER: 35,729
REFERENCE/DOCKET NUMBER: 025.6 US (IM1-028CP2)
TELEPHONE: (617) 466-6000
TELEFAX: (617) 466-6040
INFORMATION FOR SEQ ID NO: 261:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal

SQ SEQUENCE 30 AA: 3375 MW: 4852 CN:

Query Match 46.9% Score 97: DB 6: Length 30:

Best Local Similarity 56.7% Pred. No. 2,60e-01:

Matches 14: Conservative 4: Mismatches 3: Indels 0: Gaps 0:

DB 9 SSGKEGGNIYTKKEAFNVEK 29

QY 1 SSGKNEGNIYNNNEAFKVER 21

```

DE US-08-467-006-261 STANDARD: PRT: 30 AA
XX
AC xxxxxx
DE
DT
XX
DE
XX
XX
Sequence 261, Application US/08467006
Sequence 261, Application US/08467006
GENERAL INFORMATION:
APPLICANT: Griffith, Irwin J.
APPLICANT: Pollock, Joanne
APPLICANT: Bord, Julian F.
APPLICANT: Garman, Richard E.
APPLICANT: Kuo, Mei-Chang
APPLICANT: Yeung, Siu-mei H.
APPLICANT: Brauer, Andrew
APPLICANT: Exley, Mark A.
TITLE OF INVENTION: Allergenic Proteins And Peptides From
Japanese Cedar Pollen
NUMBER OF SEQUENCES: 261
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
STREET: 610 Lincoln St
CITY: Waltham
STATE: MA
COUNTRY: USA
ZIP: 02154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,006
FILING DATE: June 6, 1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/350,225
FILING DATE: December 6, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Jane E. Remillard
REGISTRATION NUMBER: 36,872
REFERENCE/DOCKET NUMBER: 025.6 US06 (1M1-28CP06)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 261
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Peptide
FRAGMENT TYPE: Internal
SEQUENCE 30 AA: 3375 MW: 4852 CN:
Query Match 46.9% Score 97: DB 10: Length 30:
Rest Local Similarity 66.7% Pred. No: 2.60e-31:
Matches 14: Conservative 4: Mismatches 3: Indels 0: Gaps 0:
DB 9 SSGKYEKGNIYTKKEAFNVEK 29
QY 1 SSGKNEGINYNNEAFKVER 21
RESULT 12
ID US-08-467-697-261 STANDARD: PRT: 30 AA
XX
AC xxxxxx
DE
DT
XX
Sequence 261, Application US/08468940
GENERAL INFORMATION:
APPLICANT: Griffith, Irwin J.
APPLICANT: Pollock, Joanne

```

```

DE Sequence 261, Application US/08467697
XX
AC xxxxxx
DE
DT
XX
DE
XX
XX
Sequence 261, Application US/08467697
Sequence 261, Application US/08467697
GENERAL INFORMATION:
APPLICANT: Griffith, Irwin J.
APPLICANT: Pollock, Joanne
APPLICANT: Bord, Julian F.
APPLICANT: Garman, Richard E.
APPLICANT: Kuo, Mei-Chang
APPLICANT: Yeung, Siu-mei H.
APPLICANT: Brauer, Andrew
APPLICANT: Exley, Mark A.
APPLICANT: Powers, Steven P.
TITLE OF INVENTION: Allergenic Proteins And Peptides From
Japanese Cedar Pollen
NUMBER OF SEQUENCES: 261
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
STREET: 610 Lincoln St
CITY: Waltham
STATE: MA
COUNTRY: USA
ZIP: 02154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,697
FILING DATE: June 6, 1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/350,225
FILING DATE: December 6, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Jane E. Remillard
REGISTRATION NUMBER: 36,872
REFERENCE/DOCKET NUMBER: 025.6 US04 (1M1-287104)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 261
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Peptide
FRAGMENT TYPE: Internal
SEQUENCE 30 AA: 3375 MW: 4852 CN:
Query Match 46.9% Score 97: DB 10: Length 30:
Rest Local Similarity 66.7% Pred. No: 2.60e-31:
Matches 14: Conservative 4: Mismatches 3: Indels 0: Gaps 0:
DB 9 SSGKYEKGNIYTKKEAFNVEK 29
QY 1 SSGKNEGINYNNEAFKVER 21
RESULT 13
ID US-08-466-940-261 STANDARD: PRT: 30 AA
XX
AC xxxxxx
DE
DT
XX
Sequence 261, Application US/08468940
GENERAL INFORMATION:
APPLICANT: Griffith, Irwin J.
APPLICANT: Pollock, Joanne

```


CC APPLICANT: Bond, Julian F.
 CC APPLICANT: Garman, Richard D.
 CC APPLICANT: Kuo, Mei-Chang
 CC APPLICANT: Yeung, Siu-mei H.
 CC APPLICANT: Brauer, Andrew
 CC APPLICANT: Exley, Mark A.
 CC APPLICANT: Powers, Steven P.
 CC TITLE OF INVENTION: Antigenic Proteins And Peptides From
 CC TITLE OF INVENTION: Japanese Cedar Pollen
 CC NUMBER OF SEQUENCES: 26
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
 CC STREET: 610 Lincoln St
 CC CITY: Waltham
 CC STATE: MA
 CC COUNTRY: USA
 CC ZIP: 02154
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patent In Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/98/468,940
 CC FILING DATE:
 CC CLASSIFICATION: 424
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/350,225
 CC FILING DATE: December 6, 1994
 CC APPLICATION NUMBER: 08/226,248
 CC FILING DATE: April 8, 1994
 CC APPLICATION NUMBER: 07/938,990
 CC FILING DATE: September 1, 1992
 CC APPLICATION NUMBER: PCT/US93/00139
 CC FILING DATE: January 15, 1993
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Darlene A. Vanstone
 CC REGISTRATION NUMBER: 35,729
 CC REFERENCE/DOCKET NUMBER: 025.6 US (IM-028CP2)
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (517) 466-6000
 CC TELEFAX: (617) 466-6040
 CC INFORMATION FOR SEQ ID NO: 261:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 30 amino acids
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: Peptide
 CC FRAGMENT TYPE: Synthetic
 CC SEQUENCE 30 AA: 339 MW: 4852 GN:

Query Match 45.9% Score 97 DB 10 Length 30
 Best Local Similarity 65.7% Pred. No. 2,60e-01
 Matches 14: Conservative 4 Mismatches 3 Indels 0 Gaps 0

DB 9 SSGKYEAGNIYTKKEAFNVEK 29
 ||||| ||||| ||||| |||||
 QY 1 SSGKNEGTNIYNNNEAFKVE 21

RESULT 14
 ID US-09-468-940-57 STANDARD: PRT: 20 AA
 XX xxxxxx
 XX
 XX
 XX
 XX

Sequence 57, Application US/38468940
 Sequence 57, Application US/08468940
 GENERAL INFORMATION:
 APPLICANT: Griffith, Irwin J.
 APPLICANT: Pollock, Joanne

CC APPLICANT: Bond, Julian F.
 CC APPLICANT: Garman, Richard D.
 CC APPLICANT: Kuo, Mei-Chang
 CC APPLICANT: Yeung, Siu-mei H.
 CC APPLICANT: Brauer, Andrew
 CC APPLICANT: Exley, Mark A.
 CC APPLICANT: Powers, Steven P.
 CC TITLE OF INVENTION: Antigenic Proteins And Peptides From
 CC TITLE OF INVENTION: Japanese Cedar Pollen
 CC NUMBER OF SEQUENCES: 26
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
 CC STREET: 610 Lincoln St
 CC CITY: Waltham
 CC STATE: MA
 CC COUNTRY: USA
 CC ZIP: 02154
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patent In Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/98/468,940
 CC FILING DATE:
 CC CLASSIFICATION: 424
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/350,225
 CC FILING DATE: December 6, 1994
 CC APPLICATION NUMBER: 08/226,248
 CC FILING DATE: April 8, 1994
 CC APPLICATION NUMBER: 07/938,990
 CC FILING DATE: September 1, 1992
 CC APPLICATION NUMBER: PCT/US93/00139
 CC FILING DATE: January 15, 1993
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Darlene A. Vanstone
 CC REGISTRATION NUMBER: 35,729
 CC REFERENCE/DOCKET NUMBER: 025.6 US (IM-028CP2)
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (517) 466-6000
 CC TELEFAX: (617) 466-6040
 CC INFORMATION FOR SEQ ID NO: 57:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 20 amino acids
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: Peptide
 CC FRAGMENT TYPE: Synthetic
 CC SEQUENCE 20 AA: 221 MW: 2114 GN

Query Match 44.9% Score 93 DB 10 Length 20
 Best Local Similarity 70.0% Pred. No. 6,84e-01
 Matches 14: Conservative 2 Mismatches 3 Indels 0 Gaps 0

DB 1 SSGKYEAGNIYTKKEAFNVE 20
 ||||| ||||| ||||| |||||
 QY 1 SSGKNEGTNIYNNNEAFKVE 20

RESULT 15
 ID US-07-938-990A-57 STANDARD: PRT: 20 AA
 XX xxxxxx
 XX
 XX
 XX
 XX

Sequence 57, Application US/07938990A
 Sequence 57, Application US/07938990A
 GENERAL INFORMATION:
 APPLICANT: Griffith, Irwin J.
 APPLICANT: Pollock, Joanne

CC APPLICANT: Bond Cullen
 CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
 CC TITLE OF INVENTION: Japanese Cedar Pollen
 CC NUMBER OF SEQUENCES: 70
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Lahive & Cockfield
 CC STREET: Sixty State Street
 CC CITY: Boston
 CC STATE: MA
 CC COUNTRY: USA
 CC ZIP: 02109
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patent Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: JS/07/935,990A
 CC FILING DATE: 19920901
 CC CLASSIFICATION: 435
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 07/730,452
 CC FILING DATE: July 15, 1991
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 07/729,134
 CC FILING DATE: July 10, 1991
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Amy E. Macdragouras
 CC REGISTRATION NUMBER: 36,207
 CC REFERENCE/DOCKET NUMBER: IPC-025CC (INT-028)
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (617) 227-7400
 CC TELEFAX: (617) 227-5941
 CC INFORMATION FOR SEQ ID NO: 57:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 20 amino acids
 CC TYPE: AMINO ACID
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: peptide
 CC FRAGMENT TYPE: internal
 CC SEQUENCE 20 AA: 222 MW: 2174 CN:

Query Match 48.94 Score 93 DB 43 Length 200
 Best Local Similarity 70.041 Pctd No. 6,846-91
 Matches 14: Conservative 4 Mismatches 3 Gaps 0

DB 1 SSGKNGGNTYKKEAFNV 20
 11111111111111111111
 GY 1 SSGKNGGNTYKKEAFNV 20

Search completed: Mon Jun 19 16:23:33 2000
 Job time : 31 secs

 W E E R E A
 (TM)

Release 3.1A John F. Collins, BioComputing Research Unit.
 Copyright (C) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
 Run on: Mon Jun 19 16:02:46 2000; MasPar time 5.20 Seconds
 Tabular output not generated. 281.442 Million cell updates/sec

Title: >US-09-142-524A-5
 Description: (1-31) from US09142524A.pep
 Perfect Score: 207
 Sequence: 1 SSGKNEGNTIYNNNEAFKVERRF:KRVSNVI 31

Scoring table: PAM 150
 Gap 11

Searched: 142080 seqs. 47172406 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: pir62
 1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 42.523; Variance 57.974; scale 0.561

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed.
 and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DP | ID | Description | Pred. No. |
|------------|-------|-------|--------|----|--------|------------------------|-----------|
| 1 | 93 | 44.9 | 374 | 2 | JC2124 | major allergen Cry 1 | 2.65e-03 |
| 2 | 93 | 44.9 | 374 | 2 | JC2123 | major allergen Cry 1 | 2.65e-03 |
| 3 | 82 | 39.6 | 344 | 2 | A47025 | abovative phage resist | 1.54e-01 |
| 4 | 82 | 39.6 | 344 | 2 | S27387 | abC protein - Lactoc | 1.54e-01 |
| 5 | 80 | 38.6 | 1101 | 2 | S66730 | hypothetical protein | 3.14e-01 |
| 6 | 78 | 37.7 | 1308 | 2 | E71622 | probable membrane ass | 6.32e-01 |
| 7 | 75 | 36.2 | 2380 | 2 | E71634 | hypothetical protein | 1.77e-00 |
| 8 | 72 | 34.5 | 256 | 2 | A32017 | beta-lactamase (EC 3. | 4.84e+00 |
| 9 | 72 | 34.8 | 260 | 1 | S11984 | fapR protein - Escher | 4.84e+00 |
| 10 | 72 | 34.8 | 317 | 2 | B72216 | endoglucanase - Ther | 4.84e+00 |
| 11 | 72 | 34.8 | 1355 | 2 | S51995 | probable ATPase (EC 3 | 4.84e+00 |
| 12 | 71 | 34.3 | 257 | 1 | PNBS25 | beta-lactamase (EC 3. | 6.72e+00 |
| 13 | 71 | 34.3 | 257 | 1 | F50727 | beta-lactamase (EC 3. | 6.72e+00 |
| 14 | 70 | 33.8 | 259 | 2 | F71810 | sensory organ homeobo | 9.31e+00 |
| 15 | 70 | 33.8 | 641 | 2 | F71810 | type II: DNA modifica | 9.31e+00 |
| 16 | 59 | 33.3 | 323 | 1 | B64330 | glycosyl transferase | 1.29e-01 |
| 17 | 59 | 33.3 | 566 | 2 | S17477 | hypothetical protein | 1.29e-01 |
| 18 | 59 | 33.3 | 1356 | 2 | S51389 | ROM2 protein - yeast | 1.29e-01 |
| 19 | 58 | 32.9 | 453 | 2 | C71518 | probable permease - C | 1.77e-01 |
| 20 | 58 | 32.9 | 610 | 2 | H71612 | asparagine--tRNA liga | 1.77e-01 |
| 21 | 58 | 32.9 | 623 | 2 | A45050 | transketolase (EC 2.2 | 1.77e-01 |
| 22 | 58 | 32.9 | 626 | 2 | A42891 | beta-galactosidase (E | 1.77e-01 |
| 23 | 58 | 32.9 | 929 | 2 | I51027 | type XI: collagen alp | 1.77e-01 |

24 67 32.4 158 2 F65103 PTS system, n-acetyly
 25 67 32.4 168 2 PN0680 nitrogenase (EC 1.18.
 26 67 32.4 275 2 A69413 conserved hypothetical
 27 67 32.4 440 2 T14735 probable serine/threo
 28 67 32.4 440 2 T14735 probable serine/threo
 29 67 32.4 440 2 T14735 probable serine/threo
 30 66 31.9 1234 2 B36186 I factor 2 (transposo
 31 66 31.9 169 2 S77820 probable exinuclease
 32 66 31.9 298 2 T25906 hypothetical protein
 33 66 31.9 353 2 D69001 conserved hypothetical
 34 66 31.9 393 2 A75205 bacteriochlorophyll s
 35 66 31.9 597 1 NBH3C4 C4B-binding protein
 36 66 31.9 956 2 B71250 valine--tRNA ligase
 37 66 31.9 7962 2 T36345 elastic titin - human
 38 65 31.4 149 2 S13450 hemoglobin - southern
 39 65 31.4 281 2 PC5305 methionine--tRNA liga
 40 65 31.4 841 2 A41254 protein-tyrosine-phos
 41 65 31.4 1266 2 I59314 isoleucine--tRNA liga
 42 65 31.4 3433 1 GNWVKV genome polyprotein -
 43 65 31.4 3434 1 GNWVKV genome polyprotein -
 44 64 30.9 235 2 T08467 aspartate racemase (E
 45 64 30.9 1661 1 DJAD12 DNA-directed DNA poly
 45 64 30.9 1214 2 S28499 probable finger prote

ALIGNMENTS

RESULT 1 JC2124 *type complete
 ENTRY major allergen Cry 1 precursor (clone pCC1-15) - Japanese
 TITLE cedar
 ORGANISM *formal_name Cryptomeria japonica *common_name Japanese cedar
 DATE 14-Jul-1994 *sequence_revision 14-Jul-1994 *text_change
 26-Aug-1999
 JC2124
 JC2123
 Score: T: Komiyama, N.; Shimizu, K.; Kusakabe, I.; Morikubo,
 K.; Kino, K.
 Biochem. Biophys. Res. Commun. (1994) 199:619-625
 Cloning and sequencing of cDNA coding for Cry 1, a major
 allergen of Japanese cedar pollen.
 *cross-references MIM:94183234
 *molecule_type mRNA
 *residues 1-374 *label: SGN
 *cross-references GB:D26545; NID:9493633; PID:D1006087; PID:9493634
 *experimental_source pollen
 *note the authors described carbohydrate binding site for
 residue 276

CLASSIFICATION *superfamily position: 1050-1050
 KEYWORDS glycoprotein; pollen
 FEATURE
 1-21
 22-374
 *domain signal sequence *status predicted *label: SGN
 *product major allergen Cry 1 (clone pCC1-15) *status
 predicted *label: MA1
 158,191,293,354 *binding site carbohydrate (Asn) (covalent) *status
 predicted

SUMMARY *length 374 *molecular_weight 40702 *checksum 3692

Query Match 44.9%; Score 93; DB 2; Length 374;

Best Local Similarity 70.0%; Pred. No. 2.65e-03;

Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 332 SSGKVEGGNIYTKKEAFNVE 351

QY 1 SSGKNEGNTIYNNNEAFKVE 20

RESULT 2

ENTRY JC2123 *type complete

TITLE major allergen Cry 1 precursor (clone pCC1-2-2) - Japanese

ORGANISM *formal_name Cryptomeria japonica *common_name Japanese cedar

DATE 14-Jul-1994 *sequence_revision 14-Jul-1994 *text_change

```

ACCESSIONS      26-Aug-1999
REFERENCE        J03123; PC2065
AUTHORS         Sone, T.; Koriyama, N.; Shimizu, K.; Kasabe, T.; Morikubo,
                  K.; Kino, K.
JOURNAL         Biochem. Biophys. Res. Commun. (1994) 199:619-625
TITLE           Cloning and sequencing of cDNA coding for Cry 3A, a major
                  allergen of Japanese cedar pollen.
CROSS-REFERENCES MIMD:34183234
ACCESSION       J03123
MOLECULE_TYPE   RNA
RESIDUES        1-374 **label: SON
CROSS-REFERENCES GB:D26544; NID:q149353; PID:dl06086; PID:q494932
EXPERIMENTAL_SOURCE pollen
ACCESSION       PC2065
MOLECULE_TYPE   protein
RESIDUES        22-53;58-81;219-242;246-259;299-307;345-372 **label: SO2
CROSS-REFERENCES MIMD:34183234
TITLE           The authors described carbohydrate binding site for
                  residue 279
KEYWORDS         *superfamily: pectate lyase LAR59
FEATURE          glycoprotein: pollen
CLASSIFICATION  1-2:
                  *domain: signal sequence *status: predicted *label: SIGN
                  *product: major allergen Cry 3A (clone PC21-2-3) *status:
                  predicted *label: MAN
                  158-191;293,354 *binding_site: carbohydrate (ASP; (covalent) *status:
                  predicted
SUMMARY          *length: 374 *molecular-weight: 40645 *checksum: 2920
Query Match      44.98; Score 93; DB 2; Length 374;
Best Local Similarity 73.08; Pred. No. 2,85e-03;
Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 332 SSGKPEGNINIKKEAFNVE 351
      ||||| ||| |||||
QY 1 SSGKNEGTINYNNEAFKVE 20

RESULT 3
ENTRY 3
TITLE 3
ALTERNATE_NAMES 3
ORGANISM 3
DATE 3
ACCESSIONS 3
REFERENCE 3
AUTHORS 3
JOURNAL 3
TITLE 3
CROSS-REFERENCES MIMD:93054365
ACCESSION       A47025
STATUS          preliminary
MOLECULE_TYPE   DNA
RESIDUES        1-344 **label: DUR
CROSS-REFERENCES GB:M95956; NID:q149359; PID:q149360
GENETICS
GENE            abic
GENOME          plasmid
SUMMARY         *length: 344 *molecular-weight: 40128 *checksum: 3503
Query Match      39.68; Score 82; DB 2; Length 344;
Best Local Similarity 28.08; Pred. No. 1,54e-01;
Matches 7; Conservative 12; Mismatches 6; Indels 0; Gaps 0;

Db 271 GTNLFGEKFKIDQHVFVKPEII 295
      ||||| ||||| |||||
QY 7 GTNYYNNEAFKVERRFKRVSNVI 31

RESULT 4
ENTRY 4
TITLE 4
ALTERNATE_NAMES 4
ORGANISM 4
DATE 4
ACCESSIONS 4
REFERENCE 4
AUTHORS 4
JOURNAL 4
TITLE 4
CROSS-REFERENCES MIMD:93054365
ACCESSION       A47025
STATUS          preliminary
MOLECULE_TYPE   DNA
RESIDUES        1-344 **label: DUR
CROSS-REFERENCES GB:M95956; NID:q149359; PID:q149360
GENETICS
GENE            abic
GENOME          plasmid
SUMMARY         *length: 344 *molecular-weight: 40128 *checksum: 3503
Query Match      39.68; Score 82; DB 2; Length 344;
Best Local Similarity 28.08; Pred. No. 1,54e-01;
Matches 7; Conservative 12; Mismatches 6; Indels 0; Gaps 0;

Db 271 GTNLFGEKFKIDQHVFVKPEII 295
      ||||| ||||| |||||
QY 7 GTNYYNNEAFKVERRFKRVSNVI 31

```

```

ENTRY          S27587
TITLE          abic protein - lactococcus lactis (fragment)
ORGANISM       Lactococcus lactis
DATE           06-Jan-1995 *sequence_revision 06-Jan-1995 *text_change
              09-Sep-1997
ACCESSIONS     S27587
REFERENCE       Durmaz, E.; Higgins, D.L.; Klaenhammer, T.K.
              Submitted to the EMBL Data Library, June 1992
              Molecular characterization of a second abortive phase
              resistance gene present within Lactococcus lactis ME-
              2
ACCESSION       S27587
STATUS          preliminary
MOLECULE_TYPE   DNA
RESIDUES        1-344 **label: DUR
CROSS-REFERENCES EMBL:M95956; NID:q149359; PID:q149360
SUMMARY        *length: 344 *checksum: 3503
Query Match      39.68; Score 82; DB 2; Length 344;
Best Local Similarity 28.08; Pred. No. 1,54e-01;
Matches 7; Conservative 12; Mismatches 6; Indels 0; Gaps 0;

Db 271 GTNLFGEKFKIDQHVFVKPEII 295
      ||||| ||||| |||||
QY 7 GTNYYNNEAFKVERRFKRVSNVI 31

RESULT 5
ENTRY 5
TITLE 5
ALTERNATE_NAMES 5
ORGANISM       hypothetical protein YOL045w - yeast (Saccharomyces
              cerevisiae)
DATE           12-Jul-1996 *sequence_revision 12-Jul-1996 *text_change
              24-Sep-1999
ACCESSIONS     S66730
REFERENCE       Ansong, W.; As, V.; Rechmann, S.; Schwader, G.; Le-Dorn,
              C.; Voss, H.; Lehman, S.
              Submitted to the Protein Sequence Database, July 1996
ACCESSION       S66730
MOLECULE_TYPE   DNA
RESIDUES        1-1101 **label: ANS
CROSS-REFERENCES EMBL:Z74758; NID:q149846; PID:CAA99051; 1;
              PID:q25169; PID:q149848; MIPS:YOL045w
EXPERIMENTAL_SOURCE strain: 2859
GENETICS
MAP_POSITION     151
CLASSIFICATION   *superfamily: class: 151; DB 2; Length 1101;
KEYWORDS         Kinases: protein kinase homology
FEATURE          ATP
SUMMARY          *domain: protein kinase homology *label: KIN
              *region: protein kinase ATP-binding motif
              *length: 1101 *molecular-weight: 124371 *checksum: 1754
Query Match      38.68; Score 60; DB 2; Length 1101;
Best Local Similarity 25.98; Pred. No. 3,14e-01;
Matches 7; Conservative 11; Mismatches 8; Indels 0; Gaps 0;

Db 773 QDKSKFVKDDNFVKDEKFMRIEP 798
      ||||| ||||| |||||
QY 6 EGTNYNNEAFKVERRFKRVSNVI 31

RESULT 6
ENTRY 6
TITLE 6
ALTERNATE_NAMES 6
ORGANISM       Probable membrane associated protein PF0125C - malaria
              parasite (Plasmodium falciparum)
DATE           13-Nov-1998 *sequence_revision 13-Nov-1998 *text_change
              07-May-1999
ACCESSIONS     E71622

```

```

REFERENCE
#authors
A71600
Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.;
Aravind, L.; Koonin, E.V.; Shalloo, S.; Mason, T.; Yu, K.;
Fuji, C.; Pederson, J.; Shen, K.; Jing, J.; Aston, C.;
Lai, Z.; Schwartz, D.C.; Perlea, M.; Salzberg, S.; Zhou,
L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.;
Fraser, C.M.; Adams, M.D.; Venter, J.C.; Hoffman, S.L.;
Science (1998) 282:1126-1132
#journal
Chromosome 2 sequence of the human malaria parasite
#title
Plasmodium falciparum.
#cross-references MIM:99021743
#accession
E71604
#status
preliminary: nucleic acid sequence not shown;
translation: not shown;

##molecule_type DNA
##residues 1-1308 ##label GAR
##cross-references GB:AE001374; GB:AE001362; NID:q3845100; PID:q3845102;
TIGR:PF08125C
##experimental_source clone 307

GENETICS
#gene
PF08125C
#summary
#length 1308 #molecular-weight 155585 #checksum 5518
Query Match 37.7% Score 78; DB 2; Length 1308;
Best Local Similarity 46.2%; Pred. No. 6,12e-01;
Matches 12; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

DL 796 MNEEDFYENNETFECED:FLKREDN 821
      |||||
QY 4 KNEGTYNNNEAFKVERFIKRVSN 29

RESULT 7
ENTRY
#type complete
#title
hypothetical protein PF0870W - malaria parasite (Plasmodium
falciparum)
#organism
#formal_name Plasmodium falciparum
#date
13-Nov-1998 #sequence-revision 13-Nov-1998 #text_change
07-May-1999
#accessions
E71604
#reference
#authors
Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.;
Aravind, L.; Koonin, E.V.; Shalloo, S.; Mason, T.; Yu, K.;
Fuji, C.; Pederson, J.; Shen, K.; Jing, J.; Aston, C.;
Lai, Z.; Schwartz, D.C.; Perlea, M.; Salzberg, S.; Zhou,
L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.;
Fraser, C.M.; Adams, M.D.; Venter, J.C.; Hoffman, S.L.;
Science (1998) 282:1126-1132
#journal
Chromosome 2 sequence of the human malaria parasite
#title
Plasmodium falciparum.
#cross-references MIM:99021743
#accession
E71604
#status
preliminary: nucleic acid sequence not shown;
translation: not shown;

##molecule_type DNA
##residues 1-2380 ##label GAR
##cross-references GB:AE001421; GB:AE001362; NID:q3845293; PID:q3845297;
TIGR:PF0870W
##experimental_source clone 307

GENETICS
#gene
PF0870W
#summary
#length 2380 #molecular-weight 283618 #checksum 2523
Query Match 36.2% Score 75; DB 2; Length 2380;
Best Local Similarity 29.2%; Pred. No. 1,77e-00;
Matches 7; Conservative 13; Mismatches 3; Indels 1; Gaps 1;

DL 122: EFLNDNLKIDKFLKKN:TIMI 1244
      |||||
QY 9 NIYNNNEAFKVERFIKRVSNV 31

RESULT 8

```

```

ENTRY
#type complete
#title
beta-lactamase (EC 3.5.2.6) II precursor - Bacillus cereus
(strain 57B/6)
#organism
#formal_name Bacillus cereus
#date
07-Jun-1990 #sequence-revision 07-Jun-1990 #text_change
22-Jun-1999
#accessions
A32017
#reference
A32017
Liu, H.M.; Rene, J.J.; Shaw, P.W.;
J. Bacteriol. (1988) 170:2873-2878
#journal
Cloning, nucleotide sequence, and expression of the Bacillus
cereus 57B/6 beta-lactamase II structural gene.
#cross-references MIM:88227879
#accession
A32017
##molecule_type DNA
##residues 1-256 ##label LIM
##cross-references GB:M15530; NID:q143127; PID:AA24562.1; PID:q143128
#classification
#superfamily beta-lactamase II
#keywords
antibiotic resistance; hydrolase
#feature
1-29
domain signal sequence #status predicted #label S'GN
30-256
#product beta-lactamase II #status predicted #label MA:
#length 256 #molecular-weight 28038 #checksum 682
#summary
Query Match 34.8% Score 72; DB 2; Length 256;
Best Local Similarity 57.1%; Pred. No. 4,84e-00;
Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 99 MVEKKFKKRVTDVI 112
      |||||
QY 18 KVERFIKRVSNVI 31

RESULT 9
ENTRY
#type complete
#title
fapR protein - Escherichia coli
#organism
#formal_name Esch. richia coli
#date
10-Sep-1999 #sequence-revision 10-Sep-1999 #text_change
10-Sep-1999
#accessions
S11984
#reference
S11984
Klaesen, P.; de la, F.K.;
MOL. Microbiol. (1990) 4:1779-1783
#journal
Characterization of fapR, a positive regulator of expression
of the 987P operon in enterotoxigenic Escherichia coli.
#cross-references MIM:31171579
#accession
S11984
#status
preliminary
##molecule_type DNA
##residues 1-260 ##label KLA
##cross-references EMBL:X53494
#note
the sequence from Fig. 3 is inconsistent with that from
Fig. 2 in lacking 96-Lys and in having an additional
Lys after 99-Phe

GENETICS
#gene
fapR
#classification
#superfamily fapR protein
#keywords
DNA binding; transcription regulation
#summary
#length 260 #molecular-weight 30349 #checksum 5453
Query Match 34.8% Score 72; DB 1; Length 260;
Best Local Similarity 40.0%; Pred. No. 4,84e-00;
Matches 10; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

Db 83 AAESYTKRSFKRKRFVKSCSIVI 107
      |||||
QY 7 GTNYYNNNEAFKVERFIKRVSNVI 31

RESULT 10
ENTRY
#type complete
#title
endoglucanase - Thermotoga maritima (strain MSB8)
#organism
#formal_name Thermotoga maritima

```

```

DATE 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change
11-Jun-1999
ACCESSIONS 972216
REFERENCE A2200
#authors Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Guin, M.L.; Dodson,
R.J.; Haft, D.H.; Hickey, E.K.; Peterson, J.D.; Nelson,
K.E.; Ketchum, K.A.; McDonald, L.; Utterback, T.R.; Malek,
J.A.; Linhorst, K.D.; Garrett, M.M.; Stewart, A.M.; Cotton,
M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
Reidberger, J.; Sutton, G.G.; Fleischmann, R.D.; White, C.;
Salzberg, S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
#journal Nature (1999) 399:323-329
#title Evidence for lateral gene transfer between Archaea and
Bacteria from genome sequence of Thermococcus maritima.
#cross-references EMBL:9287315
#accession 972216
#status Preliminary
#molecule_type DNA
#residues 1-317 #label ARN
#cross-references GB:A001913; GR:A000512; NC:04982321; PID:04992329;
TIGR:TM1751
#experimental_source strain: MS88
GENETICS
#gene TM751
SUMMARY
#length 317 #molecular_weight 37393 #checksum 6220
Query Match 34.8%; Score 72; DB 2; Length 317;
Best Local Similarity 40.0%; Pred. No. 4,94e-00;
Matches 10; Conservative 7; Mismatches 7; Indels 1; Gaps 1;
DB 57 STRAYAPP PYKINGORVEKPVDEV: 80
QY 7 GTNLYNNNEAFKVERFKVSNVI 3;

RESULT 1;
ENTRY S51995 #type complete
#probable ApPase (EC 3.6.1.-) DRF2 - yeast (Saccharomyces
cerevisiae)
ALTERNATE_NAMES protein YAL026c
ORGANISM "normal_name Saccharomyces cerevisiae"
DATE 05-May-1995 #sequence_revision 01-Sep-1995 #text_change
17-Mar-1999
ACCESSIONS S51995; B54591; S30768
REFERENCE B54596
#authors Bussey, H.; Kaback, D.B.; Zhong, W.; Vo, D.T.; Clark, M.W.;
Forlin, N.; Hahn, J.; Guilletto, P.F.; Keng, T.; Barton,
A.B.; Su, Y.; Davies, C.K.; Storms, R.K.
#submission Submitted to the EMBL Data Library August 1994
#description The sequence of chromosome I of Saccharomyces cerevisiae
#accession S51995
#molecule_type DNA
#residues 1-1355 #label BUS
#cross-references EMBL:012980; NID:q132653; PID:0595560; MIPS:YA-025c
AS4591
#authors Ripmaster, T.L.; Vaughn, G.P.; Woolford Jr., J.L.
#journal Mol. Cell. Biol. (1993) 13:7901-7912
#title DRF1 to DRF7, novel genes required for ribosome assembly and
function in Saccharomyces cerevisiae.
#cross-references EMBL:94067151
#accession B54591
#molecule_type DNA
#residues 1-890; NT:893-1355 #label RIP
#cross-references EMBL:L01795
REFERENCE S30768
#authors Ripmaster, T.L.
#submission Submitted to the EMBL Data Library, September 1992
#description Identification of genes involved in ribosome assembly.
#accession S30768
#molecule_type DNA
#residues 1-1355 #label RI2
#cross-references EMBL:L01795; NID:q171113; PID:a:71114
GENETICS

```

```

#gene SGD:DRS2; FUN18
#map_position 11 4; MIPS:YAL026c
FUNCTION probable hydrolase
KEYWORDS hydrolase; trans-membrane protein
FEATURE
251-267 #domain transmembrane #status Predicted #label PKA
448-454 #domain transmembrane #status Predicted #label PKA
498-514 #domain transmembrane #status Predicted #label PKA
1099-1115 #domain transmembrane #status Predicted #label PKA
1136-1152 #domain transmembrane #status Predicted #label PKA
1162-1178 #domain transmembrane #status Predicted #label PKA
1202-1218 #domain transmembrane #status Predicted #label PKA
SUMMARY
#length 1355 #molecular_weight 154843 #checksum 7124
Query Match 34.8%; Score 72; DB 2; Length 1355;
Best Local Similarity 33.3%; Pred. No. 4,54e-00;
Matches 10; Conservative 14; Mismatches 5; Indels 3; Gaps 3;
DB 280 NNSTAEIFSEAHDDF-VEKRWII:PVGII 308
QY 4 KNEGTNIYNN-NEAFKVERFKVSNVI 3;

RESULT 12
ENTRY PNBS2S #type complete
#beta-lactamase (EC 3.5.2.6) II precursor - Bacillus sp.
ALTERNATE_NAMES penicillinase II
ORGANISM "normal_name Bacillus sp."
DATE 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change
18-Jun-1999
ACCESSIONS A24393; B24393
REFERENCE A24393
#authors Kato, C.; Kido, T.; Watanabe, K.; Horikoshi, K.
#journal J. Gen. Microbiol. (1985) 131:3317-3324
#title Nucleotide sequence of the beta-lactamase gene of
a alkalophilic Bacillus sp. strain 170.
#cross-references EMBL:86170399
#accession A24393
#molecule_type DNA
#residues 1-257 #label KAT
#cross-references GB:M15350; NID:q:42601; PID:AAA2275 1; PID:a:42602
#experimental_source strain: 170
#accession B24393
#molecule_type protein
#residues 31-48 #label KAL2
CLASSIFICATION #superfamily beta-lactamase II
KEYWORDS antibiotic resistance; hydrolase; metalloprotease
FEATURE
1-30 #domain signal sequence #status Predicted #label SIG
31-257 #product beta-lactamase II #status Predicted #label BL2
67 #active-site GLT #status Predicted
#binding-site zinc, high affinity (His, His, Cys, His)
#status Predicted
SUMMARY
#length 257 #molecular_weight 28153 #checksum 2375
Query Match 34.3%; Score 71; DB 1; Length 257;
Best Local Similarity 57.1%; Pred. No. 6,72e-00;
Matches 8; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
DB 100 MYEKKFKQKRVTDVI 113
QY 18 KVERFKVSNVI 31;

RESULT 13
ENTRY PNBSU2 #type complete
#beta-lactamase (EC 3.5.2.6) II precursor - Bacillus cereus
ALTERNATE_NAMES cephalosporinase II; penicillinase II
ORGANISM "normal_name Bacillus cereus"
DATE 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change
18-Jun-1999

```

```

ACCESSIONS      A91806; A91344; A01006
REFERENCE        Hussain, M.; Carlino, A.; Madonna, M.J.; Lampen, J.C.
                 J. Bacteriol. (1985) 164:223-229
                 #journal
                 #title
                   beta-lactamase II gene of Bacillus cereus 569/H in
                   Escherichia coli:
                   Escherichia coli:
                   #cross-references MUID:R6008056
                   #accession A91806
                   #molecule_type DNA
                   #residues 1-257 #label HUS
                   #cross-references GB:M1189; NID:q142603; PID:q142604
                   #experimental_source strain 569/H
REFERENCE        A91344
                 #authors
                   Apler, R.P.; Daniel, M.; Fleming, J.; Hermoso, J.M.; Pang,
                   C.; Waley, S.G.
                 #journal
                   FEBS Lett. (1985) 189:207-211
                 #title
                   The amino acid sequence of the zinc-requiring beta-lactamase
                   II from the bacterium Bacillus cereus 569.
                 #cross-references MUID:86005446
                 #accession A91344
                 #molecule_type protein;
                 #residues 31-183;187-210;214-257 #label AMS
                 #experimental_source strain 569/H/9
COMMENT          beta-lactamase II from Bacillus spp. and beta-lactamase L-1 from
                 Pseudomonas maltophilia are classified together as class B
                 beta-lactamases, which hydrolyze cephalosporins and penicillins.
                 beta-lactamase II binds two Zn(II) ions per molecule. Zinc at the
                 higher affinity site is necessary for activity of the enzyme.
                 Zinc at the lower affinity site (not specified) increases
                 hydrolysis of cephalosporin C but not of benzylpenicillin.
CLASSIFICATION  #superfamily beta-lactamase II
KEYWORDS         antibiotic resistance; hydrolase; metalloprotein; zinc
FEATURE          1-30
                 #domain signal sequence #status predicted #label SIG\
                 #product beta-lactamase II #status experimental #label
                 MPN
                 67
                 #active_site Glu #status predicted\
                 #binding_site zinc, hlg affinity (His, His, Cys, His)
                 #status experimental
SUMMARY          #length 257 #molecular_weight 28092 #checksum 2899
Query Match     34.8% Score 711 DB 15 Length 257
Best Local Similarity 57.1% Pred. No. 6,72e+00
Matches         8: Conservative 50 Mismatches 10 Indels 0 Gaps 0
DE              100 MVEKKKQAVEDVI 113
                 :|||:|||||
QY              15 KVERRFKRVSNV 31
                 :|||:|||||
RESULT          14
ENTRY           #type complete
TITLE          Sensory organ homeobox protein SOHO - chicken
ORGANISM       #formal_name Gallus gallus #common_name chicken
DATE           13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change
               24-Sep-1999
ACCESSIONS      I50727
REFERENCE        Deltcher, D.L.; Pekete, D.M.; Cepko, G.L.
                 J. Neurosci. (1994) 14:486-498
                 #journal
                 #title
                   Asymmetric expression of a novel homeobox gene in vertebrate
                   sensory organs.
                 #cross-references MUID:9413287
                 #accession I50727
                 #status
                   Preliminary: translated from GB/EMBL/DBJ
                 #molecule_type mRNA
                 #residues 1-259 #label DEI
                 #cross-references EMBL:U35815; NID:q1016799; PID:AAA79164.1;
                   PID:q1016800
GENETICS        S_HO-1
                 #gene
                 #superfamily unassigned homeobox proteins; homeobox homology

```

```

KEYWORDS        DNA binding; homeobox; nucleus; transcription regulation
FEATURE          128-184
                 #domain homeobox homology #label HOX
SUMMARY          #length 259 #molecular_weight 27926 #checksum 1918
Query Match     33.8% Score 701 DB 2 Length 259
Best Local Similarity 29.6% Pred. No. 9,31e+00
Matches         8: Conservative 10 Mismatches 8 Indels 1 Gaps 1
DB              124 AGCKKRTITFSQVFQLESTFDVKR 150
QY              1 SSGKNECTIYNNEAFKVERPF-1KR 26
                 :|||:|||||:|:|
RESULT          15
ENTRY           #type complete
TITLE          Type III DNA modification enzyme (methyltransferase);
                 Helicobacter pylori (strain J99)
ORGANISM       #formal_name Helicobacter pylori
                 #variety
                 #strain J99
DATE           12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change
               12-Feb-1999
ACCESSIONS      F71810
REFERENCE        Alm, R.A.; Ling, L.S.L.; Moir, D.T.; King, B.; Brown, E.D.;
                 Dolg, P.C.; Smith, D.R.; Noonan, B.; Galid, B.C.; Delongue,
                 B.L.; Carrel, G.; Tummino, P.J.; Caruso, A.;
                 Uria-Nickelsen, M.; Mills, D.M.; Ives, C.; Gibson, R.;
                 Merberg, D.; Mills, S.D.; Jiang, Q.; Taylor, D.E.; Voris,
                 G.F.; Trust, T.G.
                 #journal
                   Nature (1999) 397:176-180
                 #title
                   Genomic sequence comparison of two unrelated isolates of the
                   human gastric pathogen Helicobacter pylori.
                 #cross-references MUID:99120557
                 #accession F71810
                 #status preliminary
                 #molecule_type DNA
                 #residues 1-641 #label ARN
                 #cross-references GB:AE001563; GB:AE001439; NID:q4156027; PID:q4156030
                 #experimental_source strain J99
GENETICS
                 #gene
                 #mod_2
                 #length 641 #molecular_weight 72947 #checksum 9519
SUMMARY          Query Match 33.8% Score 701 DB 2 Length 641
                 Best Local Similarity 52.4% Pred. No. 9,31e+00
                 Matches 11: Conservative 10 Mismatches 5 Indels 4 Gaps 2
DB              440 SEGCGLENGUNAVFK-ERREI 454
QY              5 NEGTYIYNN-NEAFKVERRFI 24
                 :|||:||||:|:|
Search completed: Mon Jun 19 16:03:08 2000
Job time : 22 secs.

```


Query Match: 33.3% Score 69; DB 1: Length 566;

Best Local Similarity 39.1%; Pred. No. 4; Cile-00;

Matches 9; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

DB 158 DVHAKLEKLVVERREKRLSDIT 180

QY 9 NLYNNKFAKVERREIKVSNVI 31

RESULT 1:
ID ROM2_YEAST STANDARD: PRI: 1356 AA;

AC P51862

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DI 15-JUL-1999 (Rel. 38, Last annotation update)

DE ROM1 GRP-GRP EXCHANGE PROTEIN 2

GN ROM2 OR YLR371W OR YLR371.3

OS Saccharomyces cerevisiae (Baker's yeast)

OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;

CC Saccharomycetaceae; Saccharomyces

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN: S288C / AB972

RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,

FA Kavali A., Patton L., Gattung S., Greco T., Kirsten J.,

RA Kucaba T., Hallsworth K., Hawkins J., Hillier L., Chier M.,

RA Johnson D., Johnston L., Langston Y., Latreille P., Le T.,

RA Mardis E., Mezeses S., Miller N., Nhan M., Pauley A., Peluso D.,

RA Rifkin B., Riley S., Tach A., Trevisakis E., Viazati D.,

RA Wilcox L., Wolldan P., Vadin M., Wilson R., Wierston R.,

RL Submitted (JAN-1995) to the EMBL/GenBank/CCSF databases.

RN [2]

RP SEQUENCE FROM N.A.

RC CHARACTERIZATION

RX MEDLINE: 95208504

RA Ozaki K., Takaka K., Imamura H., Hihara T., Kameyama T.,

RA Nonaka H., Hirano H., Matsura Y., Takai Y.,

RT "Rom1p and Rom2p are GTP/GTP exchange proteins (GEPs) for the Rom1p

small GTP binding protein in Saccharomyces cerevisiae."

RL EMBL J. 15:2196-2207 (1996).

CC -1- FUNCTION: STIMULATES THE EXCHANGE OF ROM1 GTP-BOUND FORM INTO

GTP-BOUND FORM.

CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.

CC This SWISS-Prot entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL database.

CC The European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed, changed or deleted.

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

or send an email to license@isb-sib.ch).

CC

EMBL: U19133; AB067564

SGD: U000182; ROM2

DR PFAM: PF00619; DEP: 1

DR PFAM: PF00621; RCGF: 1

DR PFAM: PF00780; CNH: 1

KW Guanine-nucleotide releasing factor.

FT DOMAIN ?

FT DOMAIN 252 265 POLY-ASN

FT DOMAIN 349 356 POLY-HIS

FT DOMAIN 632 635 POLY-ASP

SEQUENCE 1356 AA: 152595 MW: 5FBC542114E78C92 CRC64:

Query Match 33.3% Score 69; DB 1: Length 1356;

Best Local Similarity 27.3%; Pred. No. 4; Cile-00;

Matches 6; Conservative 11; Mismatches 4; Indels 1; Gaps 1;

DB 706 HYPHAINDIYNNRRELKALID 727

QY 9 NLYNNKFAKVERREIKVSNV 29

RESULT 12

TKT_HUMAN STANDARD: PRI: 623 AA;
P29401
01-DEC-1992 (Rel. 24, Created)
01-OCT-1996 (Rel. 34, Last sequence update)
01-OCT-1996 (Rel. 34, Last annotation update)
DE TRANSEKTOLOASE (EC 2.2.1.1) (TK).
GN TKT1 OR TKT.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia.
OC Eutheria; Primates; Catarrhini; Hominoidea; Homo.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 93123263
RA McCool B.A., Plonk S.G., Martin P.K., Singleton C.K.,
RT "Cloning of human transketolase cDNAs and comparison of the
nucleotide sequence of the coding region to Weinicke-Korsakoff and
non-Weinicke-Korsakoff individuals."
RL J. Biol. Chem. 268:1397-1404 (1993).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE BRAIN;
Jung E.H., Sheu K.F.R.E., Szabo P., Glass J.P.
Submitted (SEP-1993) to the EMBL/GenBank/CCSF databases.
RN [3]
RP SEQUENCE FROM N.A.
RX TISSUE BRAIN;
Schenk G., Layfield R., Candy J.M., Dugaleby R.G., Nixon P.
Submitted (MAY-1996) to the EMBL/GenBank/CCSF databases.
RN [4]
RP SEQUENCE OF 224-623 FROM N.A. AND PARTIAL SEQUENCE.
RX TISSUE BRAIN;
MEDLINE: 92231878
RA Abedina M., Layfield R., Jones S.M., Nixon P.F., Matlack J.S.
RT "Nucleotide and predicted amino acid sequence of a cDNA clone
encoding part of human transketolase."
RL Biochem. Biophys. Res. Commun. 183:1159-1166 (1992).
CC -1- CATALYTIC ACTIVITY: SEPHETIDULOSE 7-PHOSPHATE + D-HYDROXYALDEHYDE
3-PHOSPHATE + D-RIBOSE 5-PHOSPHATE + D-XYLOULOSE 5-PHOSPHATE
CC -1- COFACTOR: THIAMINE PYR-5-PHOSPHATE
CC -1- SUBUNIT: HOMODIMER.
CC -1- DISEASE: HAS BEEN IMPLICATED IN THE LATENT GENETIC DISEASE
WEINICKE-KORSAKOFF SYNDROME (WKS) WHICH CAUSE SPECIFIC BRAIN
DAMAGE.
CC -1- SIMILARITY: BELONGS TO THE TRANSEKTOLOASE FAMILY.
CC This SWISS-Prot entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL database.
CC The European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed, changed or deleted.
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

Query Match

32.9% Score 68; DB 1: Length 623;

Best Local Similarity 50.0%; Pred. No. 5,72e-00; Matches 9; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

DB 472 ISRPENAIYNNNEQV 489
 QY 2 SORNEGNIYNNNEAFK 19

RESULT 13

10 BGALIEULA STANDARD; PRT: 626 AA.
 AC Q02603;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DI 15-JUL-1998 (Rel. 37, Last annotation update)
 DE BETA-GALACTOSIDASE LARGE SUBUNIT (EC 3.2.1.23) (LACTASE).
 GN LACT.
 OS Leucostoc lactis.
 OG Plasmid pN263.
 OC Bacteria: Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
 ON Leucostoc.
 RN [1].
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-6.
 RC STRAIN=N26309;
 RX MEDLINE: 92325034.
 RA David S., Stevens H., van Riel M., Simons G., de Vos W.M.;
 PT "Leucostoc lactis beta-galactosidase is encoded by two overlapping
 genes."
 RE J. Bacteriol. 174:4475-4481(1992).
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF TERMINAL, NON-REDUCING BETA-D-
 CC GALACTOSE RESIDUES IN BETA-D-GALACTOSIDES.
 CC -1- SUBUNIT: HETERODIMER OF A LARGE AND A SMALL SUBUNIT.
 CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF GLYCOSYL HYDROLASES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M92281; AAA25267.1;
 CC PIR: A42691; A42691.
 CC HSP: P00722; LBL.
 CC PRINTS: P00132; GHYDLASE2.
 CC PROSITE: PS00719; GLYCOSYLHYDROL_F2_1; 1.
 CC PRAM: PF00703; Glyco_Hydro_2; 1.
 CC Hydrolase: Glycosidase: Plasmid.
 CC ACT_SITE 318 318 HAS AN EFFECT ON THERMOSTABILITY
 CC (POTENTIAL).
 CC ACT_SITE 456 466 PROTON DONOR (BY SIMILARITY).
 CC ACT_SITE 534 534 NUCLEOPHILE (BY SIMILARITY).
 CC SEQUENCE 626 AA: 72113 MW: E95EA1548992CF7D CRC64:
 Query Match 32.9%; Score 68; DB 1; Length 626;
 Best Local Similarity 38.9%; Pred. No. 5,72e-00;
 Matches 7; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

DB 28 YHDTAEKTSRRIKSLN 45
 QY 11 YNNNEAFKVRERIKRVS 28

RESULT 14

ID CALICNOTVI STANDARD; PRT: 929 AA.
 AC Q91145;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DI 15-JUL-1998 (Rel. 36, Last annotation update)
 DE COLLAGEN ALPHA 1(XII) CHAIN (FRAGMENT).
 OS Notophthalmus viridescens (Eastern newt) (Triturus viridescens).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;

Batrachia; Caudata; Salamandridae; Salamandridae; Notophthalmus.

CC [1].
 RN RP SEQUENCE FROM N.A.
 RX MEDLINE: 95246925.
 RA Wei Y., Yang E.V., Klatt K.P., Tassava R.A.;
 RT "Monoclonal antibody M12 identifies the urodele alpha 1 chain of type
 RT XII collagen, a developmentally regulated extracellular matrix
 RT protein in regenerating newt limbs."
 RL Dev. Biol. 168:503-513(1995).
 CC -1- FUNCTION: TYPE XII COLLAGEN INTERACTS WITH TYPE I COLLAGEN-
 CC CONTAINING FIBRILS. THE COL1 DOMAIN COULD BE ASSOCIATED WITH THE
 CC SURFACE OF THE FIBRIL. AND THE COL2 AND NC3 DOMAINS MAY BE
 CC LOCALIZED IN THE PERICELLULAR MATRIX (BY SIMILARITY). COULD PLAY
 CC A DEVELOPMENTAL ROLE IN REGENERATION.
 CC -1- SUBUNIT: TRIMER OF IDENTICAL CHAINS EACH CONTAINING 190 KD OF
 CC NONTRIPLE-HELICAL SEQUENCES (BY SIMILARITY).
 CC -1- DEVELOPMENTAL STAGE: EXPRESSION STARTS AT 3 DAYS AFTER AMPUTATION
 CC IN CELLS OF THE BASAL LAYER OF THE WOUND EPITHELIUM. AT DAY 10,
 CC EXPRESSION IS FOUND IN BOTH THE BASAL WOUND EPITHELIAL CELLS AND
 CC THE DISTAL MESENCHYME CELLS. AT MID-BOD AND LATE-BOD BLASTOMA
 CC STAGES, WOUND EPITHELIUM EXPRESSION HAS DECREASED. WHEREAS THE
 CC MESENCHYME REMAINS STRONGLY ACTIVE IN TRANSCRIPTION AND SHOWS A
 CC TENDENCY TOWARD DISTAL REGONALIZATION. CONDENSING CARTILAGE SHOWS
 CC NO SIGNAL. FINALLY, AT THE LATE DUCT STAGE, EXPRESSION BECOMES
 CC LARGELY RESTRICTED TO THE PERICHONDRIUM.
 CC -1- PTM: THE TRIPLE-HELICAL TAIL IS STABILIZED BY DISULFIDE BONDS AT
 CC EACH END (BY SIMILARITY).
 CC -1- PTM: PROLINES AT THE THIRTY POSITION OF THE TRIPEPTIDE REPEATING
 CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE FAMILY OF FIBRIL-ASSOCIATED COLLAGENS
 CC WITH INTERRUPTED HELICES (FACIT).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U19494; AA880217.1;
 CC HSP: P11276; IMEN.
 CC PRAM: PF00041; 5-3; 7.
 CC PRAM: PF00092; vva; 1.
 CC Extracellular matrix: Connective tissue. Repeat. Cell adhesion.
 CC Collagen: Glycoprotein.
 CC NON_TER 1 52
 CC DOMAIN 51 52
 CC DOMAIN 51 154
 CC DOMAIN 155 245
 CC DOMAIN 246 338
 CC DOMAIN 339 432
 CC DOMAIN 433 519
 CC DOMAIN 520 612
 CC DOMAIN 613 817
 CC DOMAIN 818 907
 CC DOMAIN 908 929
 CC BINDING 231 231
 CC BINDING 324 324
 CC BINDING 415 415
 CC CARBOHYD 98 98
 CC NON_TER 929 929
 CC SEQUENCE 929 AA: AE5D7485254FD954 CRC64:
 Query Match 32.9%; Score 68; DB 1; Length 929;
 Best Local Similarity 53.3%; Pred. No. 5,72e-00;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

DB 649 FKIVRNFIISRVVEVF 663
 QY 17 FKVERFIKRVSNVI 31

DE POLLEN MAJOR ALLERGEN 1-2.
 CS Juniperus ashei (Oak white cedar).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC euphyllophytes; Spermatophyta; Coniferopsida; Coniferales;
 CC Taxodiaceae; Juniperus.
 RN
 RA SEQUENCE FROM N.A.
 RA MIDORO-HORIUTTI T.M., GILDEBLUM R.M., KUKUSKY A., WOOD T.G.,
 RA BROOKS E.G.
 RI "Molecular cloning of mountain cedar (Juniperus ashei) pollen major
 RI allergen, Jun 1."
 RL Submitted (Nov-1998) to the EMBL/GenBank/DBJ databases.
 DK EMBL: AF106663; AAO03629.1;
 DK EMBL: AF106662; AAO03628.1;
 DR MENDEL: 36544; Jundis:1086;36544.
 DR MENDEL: 36544; Jundis:1086;36545.
 SQ SEQUENCE 357 AA; 39824 MW; 4C2DB630 CRC32;

Query Match 52.2%; Score 108; DB 10; Length 367;
 Best Local Similarity 85.0%; Pred. No. 2.30e-96;
 Matches 17; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

DB 332 SSGKTEINYNNSNEAFVE 351

QY 1 NSGRNEGNIYNNNEAFVE 20

RESULT 3

ID 096129 PRELIMINARY: PRT: 1308 AA.

AC 096129

DT 01-MAY-1999 (TRENBLrel. 10, Created)

DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)

DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)

DE PREDICTED MEMBRANE ASSOCIATED PROTEIN.

GN PBO129C.

OS Plasmodium falciparum.

CC Eukaryota; Alveolata; Apicomplexa; Haemsporidia; Plasmodium.

RN [1]

RP MEDLINE: 9902174.

RA KOONER M.J., TETTELIN R., CARUCCI D.J., CUMMINGS L.M., ARAVIND L.,

RA KOONER D.V., SHALOM S., WASON T., YU K., FUJII C., PETERSON J.,

RA SHEN K., GINE J., ASTON J., CAI Z., SCHWARTZ D., PERIEA M.,

RA SALZBERG S., ZHOU L., SUTIN S.G., CLAYTON K., WHITE G., SMITH R.D.,

RA FRASER C.M., ADAMS M., VENTER A.J., BERNAN S.D.

RT "Chromosome 2 sequence of the human malaria parasite Plasmodium

RT falciparum."

RL Science 282:1126-1132(1998).

DR EMBL: AE001374; AAC71815.1;

SQ SEQUENCE 1308 AA; 155585 MW; 58DA1924 CRC32;

Query Match 37.7%; Score 78; DB 5; Length 1308;

Best Local Similarity 46.2%; Pred. No. 3.97e-01;

Matches 12; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

DB 796 MNEEDFVWNETFECEDEKREDN 821

QY 4 KNEGINYNNEAFKVERFKRVN 29

RESULT 4

ID 092V64 PRELIMINARY: PRT: 204 AA.

AC 092V64

DT 01-MAY-1999 (TRENBLrel. 10, Created)

DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)

DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)

DE T2P1.13 PROTEIN.

GN T2P1.13.

OS Arabidopsis thaliana (Mouse-ear cress).

CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

CC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;

CC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;

CC Arabidopsis.

RN
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, COLUMBIA;
 RA FEDERSPIEL N.A., PALM C.J., CONWAY A.B., CONN L., HANSEN N.F.,
 RA ALTAFI H., APAUJO R., HUIZAR L., ROWLEY D., BUEHLER E., JENN P.,
 RA GONZALEZ A., KREMEZEISKAIA I., KIM C., LENZ C., LI J., LIU S.,
 RA LUCOS S., SCHWARTZ J., SHINN P., TORIUMI M., VYSOTSKAIA V.S.,
 RA WALKER M., YC G., ECKER J., THEOLOGIS A., DAVIS R.M.,
 RC Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC005508; AAD14500.1;
 DR HSSP: P20752; 1CLH.
 SQ SEQUENCE 204 AA; 24097 MW; 3C2F9CEB CRC32;

Query Match 36.7%; Score 16; DB 10; Length 204;

Best Local Similarity 40.0%; Pred. No. 8.27e-01;

Matches 8; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

DB 35 YNTNHFFRVKGFVAQVADV 54

QY 11 YNNNEAFKVERFKRVSNV 30

RESULT 5

ID 092GL9 PRELIMINARY: PRT: 360 AA.

AC 092GL9

DT 01-MAY-1999 (TRENBLrel. 10, Created)

DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)

DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)

DE HYPOTHETICAL 41.5 KD PROTEIN.

CS Leptospira borgpetersenii.

CC Bacteria; Spirochaetales; Leptospiraceae; Leptospira

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-L171;

RA KALAMBAHETI T., BULACH D.M., RAJAKUMAR K., ADLER B.

RT "Genetic organization of the lipopolysaccharide O-antigen biosynthetic

RT locus of Leptospira borgpetersenii: Serovar Hardjohovis."

RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF078135; AAD12950.1;

KW Hypothetical protein.

SQ SEQUENCE 360 AA; 41546 MW; 622D1EE3 CRC32;

Query Match 36.7%; Score 16; DB 10; Length 360;

Best Local Similarity 40.0%; Pred. No. 8.27e-01;

Matches 8; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

DB 144 FNLNFFFKRVJSVI 136

QY 17 FKVERFKRVSNVI 31

RESULT 6

ID 087939 PRELIMINARY: PRT: 548 AA.

AC 087939

DT 01-NOV-1998 (TRENBLrel. 08, Created)

DT 01-NOV-1998 (TRENBLrel. 08, Last sequence update)

DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)

DE TDIS.

GN TDIS.

OS Thauera aromatica.

CC Bacteria; Proteobacteria; beta subdivision; Rhodocyclus group;

CC Thauera.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-K172;

RX MEDLINE: 96294052.

RA LEUTHNER B., LEUTHNER C., SCHULTZ H., HOERTH P., HAEHNEL W.,

RA SCHULTZ E., SCHAEGER H., REIDER J.

RT "Biochemical and genetic characterization of benzylsuccinate synthase

RT from Thauera aromatica: a new glycyl radical enzyme catalysing the

RT first step in anaerobic toluene metabolism."

RL Mol. Microbiol. 28:625-629(1998).

DR EMBL: AJ001848; CAA05048.1;

RA STEWART A.M., COTTON M.D., PRATT M.S., PHILLIPS C.A., RICHARDSON D.,
RA HEIDELBERG J., SUTTON G.G., FLEISCHMANN R.D., WHITE C., SALZBERG S.L.,
RA SMITH H.O., VENTER J.C., FRASER C.M.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AE001813; AAD36825.1;
SQ SEQUENCE 317 AA: 37383 MW: 38FALL08 CRC32.

Query Match 34.88; Score 72; DB 2; Length 317
Best Local Similarity 40.38; Pred. No. 3,45e-00;
Matches 10; Conservative 7; Mismatches 7; Indels 1; Gaps 1;

DB 57 STAYAFP-PYKIMDFKRVDEVI 80
QY 7 GTIYNNEAFKVERFKRVSNV 31

RESULT 9
ID Q9XPX4 PRELIMINARY; PRT: 360 AA.
AC Q9XPX4;
DI 01-NOV-1999 (TrEMBLrel. 12, Created)
DI 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DI 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE MATURASE (FRAGMENT);
GN MATK.
OS Zizyphus obtusifolia.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta.
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC core eudicots; Rosidae; eurosids I; Rosales; Rhmannaceae; Ziziphus
[1].
RN SEQUENCE FROM N.A.
RC STRAIN-0622120B;
RA HARDIG T.M., SOLTIS P.S., SOLTIS D.E.;
RT "Diversification of the North American shrub genus *Ceanothus*
[Rhmannaceae]: conflicting phylogenies from nuclear ribosomal rDNA and
RT chloroplast DNA";
RL AM. J. Bot. 0:0-0(1999).
KW Chloroplast.
FT NON-TER 360 360
FT NON-TER
SQ SEQUENCE 360 AA: 42765 MW: 222A2B54 CRC32.

Query Match 34.38; Score 71; DB 8; Length 360
Best Local Similarity 35.38; Pred. No. 4,91e-00;
Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

DB 35 YERKSSPRKVERFKRV 62
QY 11 YNNNEAFKVERFKRV 47

RESULT 10
ID Q9XPX3 PRELIMINARY; PRT: 372 AA.
AC Q9XPX3;
DI 01-NOV-1999 (TrEMBLrel. 12, Created)
DI 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DI 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE MATURASE (FRAGMENT);
GN MATK.
OS Adolphia californica.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC core eudicots; Rosidae; eurosids I; Rosales; Rhmannaceae; Adolphia
[1].
RN SEQUENCE FROM N.A.
RC STRAIN-ADOLCAL;
RA HARDIG T.M., SOLTIS P.S., SOLTIS D.E.;
RT "Diversification of the North American shrub genus *Ceanothus*
[Rhmannaceae]: conflicting phylogenies from nuclear ribosomal DNA and
RT chloroplast DNA";
RL AM. J. Bot. 0:0-0(1999).

RA STEWART A.M., COTTON M.D., PRATT M.S., PHILLIPS C.A., RICHARDSON D.,
RA HEIDELBERG J., SUTTON G.G., FLEISCHMANN R.D., WHITE C., SALZBERG S.L.,
RA SMITH H.O., VENTER J.C., FRASER C.M.;
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AE001813; AAD36825.1;
SQ SEQUENCE 317 AA: 37383 MW: 38FALL08 CRC32.

Query Match 34.88; Score 72; DB 2; Length 317
Best Local Similarity 40.38; Pred. No. 3,45e-00;
Matches 10; Conservative 7; Mismatches 7; Indels 1; Gaps 1;

DB 57 STAYAFP-PYKIMDFKRVDEVI 80
QY 7 GTIYNNEAFKVERFKRVSNV 31

RESULT 9
ID Q9XPX4 PRELIMINARY; PRT: 360 AA.
AC Q9XPX4;
DI 01-NOV-1999 (TrEMBLrel. 12, Created)
DI 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DI 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE MATURASE (FRAGMENT);
GN MATK.
OS Zizyphus obtusifolia.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta.
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC core eudicots; Rosidae; eurosids I; Rosales; Rhmannaceae; Ziziphus
[1].
RN SEQUENCE FROM N.A.
RC STRAIN-0622120B;
RA HARDIG T.M., SOLTIS P.S., SOLTIS D.E.;
RT "Diversification of the North American shrub genus *Ceanothus*
[Rhmannaceae]: conflicting phylogenies from nuclear ribosomal rDNA and
RT chloroplast DNA";
RL AM. J. Bot. 0:0-0(1999).
KW Chloroplast.
FT NON-TER 360 360
FT NON-TER
SQ SEQUENCE 360 AA: 42765 MW: 222A2B54 CRC32.

Query Match 34.38; Score 71; DB 8; Length 360
Best Local Similarity 35.38; Pred. No. 4,91e-00;
Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

DB 35 YERKSSPRKVERFKRV 62
QY 11 YNNNEAFKVERFKRV 47

RESULT 10
ID Q9XPX3 PRELIMINARY; PRT: 372 AA.
AC Q9XPX3;
DI 01-NOV-1999 (TrEMBLrel. 12, Created)
DI 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DI 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE MATURASE (FRAGMENT);
GN MATK.
OS Adolphia californica.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC core eudicots; Rosidae; eurosids I; Rosales; Rhmannaceae; Adolphia
[1].
RN SEQUENCE FROM N.A.
RC STRAIN-ADOLCAL;
RA HARDIG T.M., SOLTIS P.S., SOLTIS D.E.;
RT "Diversification of the North American shrub genus *Ceanothus*
[Rhmannaceae]: conflicting phylogenies from nuclear ribosomal DNA and
RT chloroplast DNA";
RL AM. J. Bot. 0:0-0(1999).

DR EMBL: AFO49847: AAD30358.1: -
 KW Chloroplast.
 FT NON-TER 372
 SQ SEQUENCE 372 AA: 44233 MW: F5AE7E4F CRC32:

Query Match 34.3% Score 71: DB 8: Length 372:
 Best Local Similarity 35.1% Pred. NO. 4.91e+00:
 Matches 6: Conservative 8: Mismatches 3: Indels 0: Gaps 0:

DB 48 YONKSSLRVKKRPIRM 44
 11 YNNNEAFKVEREIKRV 23

RESULT 11
 ID Q9X2X5 PRELIMINARY: PR: 372 AA.

AC Q9X2X5
 DT 01-NOV-1999 (TRENBLrel. 12, Created)
 DI 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
 DE 01-NOV-1999 (TRENBLrel. 12, Last annotation update)
 DE MATURASE (FRAGMENT)
 GN MATK
 OS Spyradium parvifolium.
 CC Chloroplast.
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 CC core eudicots; Rosidae; eurosids 1; Rosales; Rhanales; Spyradium.
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN:SPYRIPARV1
 RA HARGIS T.M., SOLIS P.S., SOLIS D.E.:
 RT "Diversification of the North American shrub genus *Ceanothus*
 RT (Rhamnaceae): conflicting phylogenies from nuclear rDNA and
 RT chloroplast DNA."
 RL Am. J. Bot. 76:0-0(1999).
 DR EMBL: AFO49849: AAD30360.1: -
 KW Chloroplast.
 FT NON-TER 372
 SQ SEQUENCE 372 AA: 44233 MW: F5AE7E4F CRC32:

Query Match 34.3% Score 71: DB 8: Length 372:
 Best Local Similarity 35.1% Pred. NO. 4.91e+00:
 Matches 6: Conservative 8: Mismatches 3: Indels 0: Gaps 0:

DB 48 YONKSSLRVKKRPIRM 44
 11 YNNNEAFKVEREIKRV 23

RESULT 12
 ID Q9Y2E5 PRELIMINARY: PR: 823 AA.

AC Q9Y2E5
 DT 01-NOV-1999 (TRENBLrel. 12, Created)
 DI 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
 DE KIAA0935 PROTEIN (FRAGMENT).
 GN KIAA0935.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 CC Eutheria; Primates; Catarrhini; Homidae; Homo.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN.
 RX MEDLINE: 99246063.
 RA NAGASE T., ISH-KAWA K., SUYAMA M., KIKUNO K., HIROSAWA M.,
 RA MIYAJIMA N., TANAKA A., KOTANI H., NOMURA N., OHARA O.:
 RT "Prediction of the coding sequences of unidentified human genes. XIII.
 RT The complete sequences of 100 new cDNA clones from brain which code
 RT for large proteins in vitro."
 RL DNA Res. 6:63-76(1999).
 DR EMBL: AB023152: BAA76779.1: -
 FT NON-TER 1
 SQ SEQUENCE 823 AA: 92710 MW: BF97CB35 CRC32:

Query Match 34.3% Score 71: DB 4: Length 823:
 Best Local Similarity 34.8% Pred. NO. 4.91e+00:
 Matches 9: Conservative 10: Mismatches 4: Indels 1: Gaps 1:

DB 598 IYSDANGYQOMRRPVSYVNSI 620
 10 IYNNNEAFKVEREIKRVSNV 31

RESULT 13
 ID Q91964 PRELIMINARY: PR: 259 AA.

AC Q91964
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DI 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
 DE 01-NOV-1999 (TRENBLrel. 12, Last annotation update)
 DE SENSORY ORGAN HOMEOBOX PROTEIN S.O.H.O.
 GN SOHO-1.
 OS Gallus gallus (Chicken).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Archosauria; Anser;
 CC Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 94132877.
 RA DEITCHER D.L., PEKETE D.M., CEPKO G.L.:
 RT "Asymmetric expression of a novel homeobox gene in vertebrate sensory
 RT organs."
 RL J. Neurosci. 14:486-498(1994).
 RN [2]

RP SEQUENCE FROM N.A.
 RA DEITCHER D.:
 RT Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (SV SIMILARITY).
 DR EMBL: J35815: AAT79164.1: -
 DR EMBL: S69380: AAC60747.1: -
 DR HSSP: P22808: INK3.
 DR PROSITE: PS00027: HOMEODOMAIN 1.
 DR PFAM: PF00046: homeobox.
 DR PRINTS: PR00024: HOMEBOX.
 KW Homeobox; DNA-binding; Nuclear protein.
 SQ SEQUENCE 259 AA: 27926 MW: 03819495 CRC32:

Query Match 33.8% Score 70: DB 13: Length 259:
 Best Local Similarity 29.6% Pred. NO. 6.99e+00:
 Matches 10: Conservative 10: Mismatches 4: Indels 1: Gaps 1:

DB 124 AGKKKTRTFKSKVFESENTHVKK 190
 1 SSKKNGSNYYNNNEAFKVEREIKRV 20

RESULT 14
 ID Q03561 PRELIMINARY: PR: 505 AA.

AC Q03561
 DT 01-JUL-1997 (TRENBLrel. 04, Created)
 DI 01-AUG-1998 (TRENBLrel. 07, Last sequence update)
 DE 01-NOV-1999 (TRENBLrel. 12, Last annotation update)
 DE MATURASE.
 GN YCF14 OR MATK.
 OS Cardenia thumbergia.
 CC Chloroplast.
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 CC core eudicots; Asteridae; euasterids 1; Gentianales; Rubiaceae;
 CC Cardenia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA ENDRESS M.E., SENNBAD B., NILSSON S., CIVEYREL L., CHASE M.W.:
 RA HUYSMA S., GRAFSTROEM E., BREMER B.:
 RL Opera Bot. Belg. 7:59-102(1996).
 DR EMBL: Z70198: CAA94087.1: -
 DR MENDE : 23592; Garth:ycf14.23592
 DR PFAM: PF01348; Intronic_matur_s2; 1.

```

KW Chloroplast
SQ SEQUENCE 505 AA: 60113 MW: 50680038 CRC32:
Query Match 33.8% Score 70: DB 8: Length 505:
Best Local Similarity 33.3% Prod. No. 5.95e+00:
Matches 6: Conservative 8: Mismatches 4: Indels 0: Gaps 0:

Dd 47 VYDNKSSRLVVKRLIR: 64
      |||:|||||:
Qy 10 LYNNNEAFKVERREIKPV 27

RESULT 15
IC Q92B21 PRELIMINARY: PRI: 545 AA.
AC Q92B23.
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DI 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE HYPOTHETICAL 62.3 Kd PROTEIN.
GN TUF1.
OS Thauera aromatica.
OC Bacteria; Proteobacteria; beta subdivision; Rhodocyclus group;
OC Thauera.
RN [1].
RP SEQUENCE FROM N.A.
PC STRAIN-T1.
EX MEDLINE: 97176397.
WA COSCHIGANO P.W., YOUNG L.Y.:
RT Identification and sequence analysis of two regulatory genes involved
PI in anaerobic toluene metabolism by strain T1.
RL Appl. Environ. Microbiol. 63:652-660(1997).
RW [2].
RP SEQUENCE FROM N.A.
RC STRAIN-T1.
RA COSCHIGANO P.W., YOUNG L.Y.:
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RW [3].
RP SEQUENCE FROM N.A.
RC STRAIN-T1.
RA COSCHIGANO P.W., YOUNG L.Y.:
RT Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: U57500; AAD2287.1; .
KW Hypothetical protein.
SQ SEQUENCE 545 AA: 62313 MW: 5494023 CRC32:
Query Match 33.8% Score 70: DB 2: Length 545:
Best Local Similarity 26.1% Prod. No. 5.95e+00:
Matches 6: Conservative 12: Mismatches 5: Indels 0: Gaps 0:

Dd 84 HFERNDIYKDERFVKDSVV 106
      |||:|||||:
Qy 9 NIYNNNEAFKVERREIKRVNVI 31

Search completed: Mon Jun 19 16:02:29 2000
Job time : 21 secs

```

W P E R E A H
***** (PW)

Release 4.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution Rights by Oxford Molecular Ltd

XPsearch_pp Protein - protein database search, using Smith-Waterman algorithm
Run on: Mon Jun 19 16:19:11 2000: MasPar time 2.63 Seconds
Tabular output not generated. 21,065 Million cell updates/sec

Title: >US-09-142-524A-14
Description: (1-9); from US09142524A.pep
Perfect Score: 57
Sequence: 1 SIXRVSNVI 9

Scoring table: PAM 150
Gap 15

Searched: 149963 seqs. 21665106 residues

Post-processing: Minimum Match 04
Listing first 45 summaries

Database: a-geneseq45
l-geneseqf

Statistics: Mean 15.426, Variance 44.241, scale 0.349

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description | Pred. No. |
|------------|-------|-------------|--------|--------|-----------------------------|-----------|
| 1 | 53 | 43.0 | 9 | W57751 | Residues 109-117 of Cry J 1 | 1.87e-01 |
| 2 | 53 | 43.0 | 15 | W57751 | Residues 109-120 of Cry J 1 | 1.87e-01 |
| 3 | 53 | 43.0 | 15 | R89231 | Japanese cedar pollen | 1.87e-01 |
| 4 | 53 | 43.0 | 20 | R45552 | Cry J 1 pollen allergen | 1.87e-01 |
| 5 | 53 | 43.0 | 20 | R45551 | Cry J 1 Japanese cedar | 1.87e-01 |
| 6 | 53 | 43.0 | 30 | W44563 | T-cell epitope peptide | 1.87e-01 |
| 7 | 53 | 43.0 | 31 | W27372 | Multi-epitope peptide | 1.87e-01 |
| 8 | 53 | 43.0 | 31 | W27373 | Multi-epitope peptide | 1.87e-01 |
| 9 | 53 | 43.0 | 31 | W27366 | Multi-epitope peptide | 1.87e-01 |
| 10 | 53 | 43.0 | 105 | W27371 | Multi-epitope peptide | 1.87e-01 |
| 11 | 53 | 43.0 | 134 | W27371 | Multi-epitope peptide | 1.87e-01 |
| 12 | 53 | 43.0 | 353 | R75339 | Japanese cedar pollen | 1.87e-01 |
| 13 | 53 | 43.0 | 353 | R61597 | Cedar pollen allergen | 1.87e-01 |
| 14 | 53 | 43.0 | 374 | R60156 | Japanese cedar pollen | 1.87e-01 |
| 15 | 53 | 43.0 | 374 | R31937 | Cry J 1 | 1.87e-01 |
| 16 | 53 | 43.0 | 374 | R82490 | Cry J 1 Japanese cedar | 1.87e-01 |
| 17 | 53 | 43.0 | 374 | R25541 | Cry J 1 pollen allergen | 1.87e-01 |
| 18 | 46 | 80.7 | 20 | R25502 | Cry J 1 Japanese cedar | 1.20e-02 |
| 19 | 46 | 80.7 | 20 | R45553 | Cry J 1 pollen allergen | 1.20e-02 |
| 20 | 44 | 77.2 | 500 | R72593 | Acromonium crysocephalum | 2.02e-02 |
| 21 | 44 | 77.2 | 566 | W86348 | Human multiple endocrine | 2.02e-02 |
| 22 | 44 | 77.2 | 610 | W29719 | Tumor suppressor prote | 2.02e-02 |
| 23 | 42 | 73.7 | 292 | R34895 | Phenylalanine Insulin | 3.35e-02 |

| | | | | | | | |
|----|----|------|------|---|--------|-------------------------|----------|
| 24 | 42 | 73.7 | 641 | 2 | W11473 | Marek's disease virus | 4.35e-02 |
| 25 | 41 | 71.9 | 171 | 1 | W20206 | H. pylori derived prot | 4.31e-02 |
| 26 | 41 | 71.9 | 266 | 1 | W98839 | H. pylori GHP 1533 pr | 4.31e-02 |
| 27 | 41 | 71.9 | 276 | 1 | W20776 | H. pylori flagella-ass | 4.31e-02 |
| 28 | 41 | 71.9 | 343 | 1 | W90011 | Expressed antigen for | 4.31e-02 |
| 29 | 41 | 71.9 | 357 | 1 | R55665 | Rat RCL7 serotonergic | 4.31e-02 |
| 30 | 41 | 71.9 | 357 | 1 | R44847 | Murine SH75A serotonin | 4.31e-02 |
| 31 | 41 | 71.9 | 367 | 1 | W52279 | Babesia microti HMN1-3 | 4.31e-02 |
| 32 | 41 | 71.9 | 367 | 1 | W56301 | Babesia microti antigen | 4.31e-02 |
| 33 | 41 | 71.9 | 401 | 1 | W93395 | S. erythraea dDB-D-6-d | 4.31e-02 |
| 34 | 41 | 71.9 | 401 | 1 | W19741 | Sugar biosynthesis enz | 4.31e-02 |
| 35 | 41 | 71.9 | 445 | 1 | W15558 | Helicobacter polyptici | 4.31e-02 |
| 36 | 41 | 71.9 | 445 | 1 | W88793 | H. pylori GHP 1252 pr | 4.31e-02 |
| 37 | 41 | 71.9 | 452 | 1 | W20598 | H. pylori proteinase | 4.31e-02 |
| 38 | 41 | 71.9 | 501 | 1 | W25094 | HPV1: mutant L1 protei | 4.31e-02 |
| 39 | 41 | 71.9 | 1429 | 1 | W93941 | Human brx protein. | 4.31e-02 |
| 40 | 40 | 70.2 | 332 | 1 | W19724 | Melanocortin-4 recepto | 5.54e-02 |
| 41 | 40 | 70.2 | 332 | 1 | W95687 | Melanocortin-4 recepto | 5.54e-02 |
| 42 | 40 | 70.2 | 332 | 1 | W92442 | Human MC4 protein. | 5.54e-02 |
| 43 | 40 | 70.2 | 360 | 1 | W98912 | RPS2 protein. | 5.54e-02 |
| 44 | 40 | 70.2 | 917 | 1 | R80132 | RPS2 disease resistant | 5.54e-02 |
| 45 | 40 | 70.2 | 1054 | 1 | W30913 | CER protein. | 5.54e-02 |

ALIGNMENTS

RESULT 1
ID W57751 standard: peptide: 9 AA:
AC W57751:
DE 17-SEP-1998 (first entry)
DE Residues 109-117 of Cry J 1
KW Cry J 1: Japanese cedar pollen antigen; allergy: immunotherapy:
KW HLA class II molecule.
CS Cryptomeria japonica.
PN WC9820402-A1
PD 22-MAY-1998:
PF 12-NOV-1997: J04329.
PR 13-NOV-1996: JP-302053.
PA (MEIP) MEIJI MILK PROD CO L
PI Daiichi K. Kino, K. Kume A, S
DR WP: 98-29781/26.
PR Peptides derived from Japan cedar pollen antigens are
PT Immunotherapeutic agents - all for allergy treatment and typing
PT HLA class II molecules in allergy sufferers
PS Claim 12: Page 26; 50pp; Japanese.
CC This sequence represents residues 109-117 of the Cry J 1 protein, and
CC is a peptide of the invention. The peptides are derived from Japanese
CC cedar pollen antigens, and are used as immunotherapeutic agents in the
CC treatment of allergy. The peptides can be used for identification, and
CC typing of the particular HLA class II molecules in an allergy sufferer,
CC and also for peptide immunotherapy of an allergy. Using these peptides
CC the immunotherapy can be targeted more specifically to the recipients
CC of the individual patient, allowing more effective treatment of an
CC allergy, including those patients for whom treatment with a conventional
CC immunotherapeutic agent is ineffective.
SQ Sequence 9 AA:

Query Match 93.0% Score 53: DB 1: length 9:
Best Local Similarity 100.0% Pred. No. 1.87e-01:
Matches 8: Conservative C: Mismatches 0: Indels 0: Gaps 0:

Db 2 IKRVSNVI 9
CY 2 IKRVSNVI 9

RESULT 2
ID W57750 standard: peptide: 15 AA:
AC W57750:
DE 17-SEP-1998 (first entry)
DE Residues 106-120 of Cry J 1:
KW Cry J 1: Japanese cedar pollen antigen; allergy: immunotherapy:
KW HLA class II molecule.

US Cryptomeria japonica.
 PN W09620402-A.
 PD 22-MAY-1998.
 PE 12-NOV-1997: J04129.
 PR 13-NOV-1996: JP-302053.
 PA (MEIP) ; MEIP: MILK PROD CO LTD.
 PI Dairiki, K, Kino K, Kume A, Sone T.
 DR WPI: 98-297517/26.
 PT Peptides derived from Japanese cedar pollen antigens are
 PI immunotherapeutic agents - useful for allergy treatment and typing
 PI HLA class II molecules in allergy sufferers
 PS Claim 12: Page 26: Supp: Japanese
 CC This sequence represents residues 106-120 of the Cry j 1 protein, and
 CC is a peptide of the invention. The peptides are derived from Japanese
 CC cedar pollen antigens, and are used as immunotherapeutic agents in the
 CC treatment of allergy. The peptides can be used for identification and
 CC typing of the particular HLA class II molecules in an allergy sufferer,
 CC and also for peptide immunotherapy of an allergy, using these peptides.
 CC the immunotherapy can be targeted more specifically to the requirements
 CC of the individual patient, allowing more effective treatment of an
 CC allergy, including those patients for whom treatment with a conventional
 CC immunotherapeutic agent is ineffective.
 SQ Sequence 15 AA.

Query Match 93.0% Score 53: DB 1: Length 15:
 Best Local Similarity 100.0% Pred. No. 1.87e-01:
 Matches 8: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Db 5 IKRVSNVI 12
 IIILIIII
 QY 2 IKRVSNVI 4
 RESULT 3
 ID R89291 standard; peptide: 15 AA.
 AC R89291:
 DE 12-MAY-1996 (first entry)
 PE Japanese cedar pollen allergen (Cry) I derived 1-cell epitope peptide.
 KW Japanese cedar pollen allergen: Cry I 1-cell epitope: peptides:
 KW prevention; treatment: cryptomeria japonica: sis.
 US Cryptomeria japonica.
 PN J0716295-A
 PD 09-MAY-1995
 PE 20-OCT-1993: J072626.
 PR 20-OCT-1993: JP-262626.
 PA (MEIP) ; MEIP: MILK PROD CO LTD.
 DR WPI: 95-203834/27.
 PT New cryptomeria pollen allergen 1-cell epitope peptide - used for
 PT prevention; treatment and investigation of Japanese cedar pollenosis
 PS Claim 5: Page 2: Supp: Japanese
 CC R75388 is the Japanese cedar pollen allergen Cry I, from which the
 CC 1-cell epitope peptides R89291-R89295 were derived. The peptides
 CC can be used for the prevention and treatment of cryptomeria
 CC pollenosis, and also for the investigation of pollenosis.
 SQ Sequence 15 AA.

Query Match 93.0% Score 53: DB 1: Length 15:
 Best Local Similarity 100.0% Pred. No. 1.87e-01:
 Matches 8: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Db 5 IKRVSNVI: 12
 IIILIIII
 QY 2 IKRVSNVI 4
 RESULT 4
 ID R45552 standard; Protein: 20 AA.
 AC R45552:
 DE 13-JUL-1994 (first entry)
 PE Cry j 1 pollen allergen peptide CJI-11.
 KW Japanese cedar: detection; allergy; treatment; diagnosis:
 KW 1-cell epitope: sensitivity.
 US Cryptomeria japonica.

PN W09401560-A.
 PD 20-JAN-1994.
 PE 15-JAN-1993: U00139.
 PR 10-JUL-1992: WO-005661.
 PA (IMMU) ; IMMUNOLOGIC PHARM CORP.
 PI Bond JF, Garman RO, Griffith LL, Kuo M, Pollock J.
 DR WPI: 94-035666/94.
 PT Antigens derived from Japanese cedar pollen allergen Cry j 1
 PI contain at least two T cell epitope(s), used to treat or diagnose
 PI allergy
 PS Claim 1: Fig 13: 137pp: English
 CC the sequence is that of an isolated peptide of the Japanese cedar
 CC pollen allergen Cry j 1 (amino acids 101-120). The peptide, CJI-11,
 CC can be used for the treatment and diagnosis of allergies associated
 CC with Japanese cedar pollen. It has enhanced therapeutic properties
 CC but reduced side effects compared to naturally occurring allergens
 SQ Sequence 20 AA.

Query Match 93.0% Score 53: DB 1: Length 20:
 Best Local Similarity 100.0% Pred. No. 1.87e-01:
 Matches 8: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Db 10 IKRVSNVI 17
 IIILIIII
 QY 2 IKRVSNVI 9
 RESULT 5
 ID R42501 standard; Protein: 20 AA.
 AC R42501:
 DE 15-APR-1996 (first entry)
 PE Cry j 1 Japanese Cedar pollen allergen peptide fragment (CJI-11).
 KW Cry j 1: Japanese cedar pollen allergen: modified; drug production:
 KW allergy: Cryptomeria japonica.
 OS Cryptomeria japonica.
 PN W09527786-A1.
 PD 19-OCT-1995.
 PE 06-APR-1995: U04249
 PR 06-APR-1994: US-226248.
 PA (IMMU) ; IMMUNOLOGIC PHARM CORP.
 PI Chen X, Evans S, Franzen BM, Kuo M, Powers SP.
 DR WPI: 95-365391/47.
 PT Modified cryptomeria japonica (Cry j) 1 peptide(s) - useful for
 PT treating allergy to Japanese cedar pollen allergen or
 PT immunologically cross reacting allergens
 PS Disclosures: Figure 2: 6pp: English
 CC Novel peptides of Cry j 1 have been modified as a part of a
 CC proteinization scheme to develop an optimised drug product for
 CC therapeutic treatment of humans suffering from allergy to Japanese
 CC cedar pollen allergen or an allergen which is immunologically cross
 CC reactive with Japanese cedar pollen allergen. Such modified peptides
 CC possess certain characteristics which render them particularly
 CC suitable for drug product formulation. Peptide fragments of Cry j 1,
 CC modified and unmodified, are given in R82491-R82525. This peptide
 CC fragment corresponds to amino acids 101-120 of the allergen mature
 CC protein.
 SQ Sequence 20 AA.

Query Match 93.0% Score 53: DB 1: Length 20:
 Best Local Similarity 100.0% Pred. No. 1.87e-01:
 Matches 8: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Db 10 IKRVSNVI 17
 IIILIIII
 QY 2 IKRVSNVI 9
 RESULT 6
 ID W44683 standard; peptide: 30 AA.
 AC W44683:

```

DI 01-MAY-1996 (first entry)
DE T-cell epitope peptide #2 of sugi pollen antigen.
KW T-cell epitope; sugi pollen antigen; sugi pollinosis.
OS Synthetic.
OS Crypthomeria japonica.
PN 21002770-A.
PD 13-JAN-1998.
PF 24-JUN-1996; 163287.
PA (DAI) : DAICEL CHEM IND LTD.
PA (MEI) : MEIJI SEIKA KALSHA LTD.
DR WPI: 97-470495/43.
PI Peptide immunotherapeutic agent to treat allergic diseases.
PT Contains multi-epitope peptide containing T cell epitope regions.
PS Claim 1; Page 4; 14pp; Japanese.
CC T-cell epitope peptides W4582-68 and their derivatives react with
CC sugi pollinosis patient peripheral blood T lymphocytes. A composition
CC prepared by combining at least 2 of the above peptides and/or their
CC derivatives is used for the prevention and treatment of sugi pollinosis.
SQ Sequence 30 AA:

Query Match 93.0% Score 53; DB 1; Length 30;
Best Local Similarity 100.0%; Pred. No. 1.87e-01;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DE 10 IKRVSNVI 17
      |||||
QY 2 IKRVSNVI 9

RESULT 7
ID W27372 standard; peptide: 31 AA.
AC W27372:
DE 24-MAR-1996 (first entry)
DE Multi-epitope peptide used as immunotherapeutic agent #4.
KW Multi-epitope peptide; immunotherapeutic agent; allergic disease.
KW T-cell epitope region; allergen; lymphocyte; immunoglobulin E.
OS Synthetic.
OS Crypthomeria japonica.
PN 21002770-A.
PD 12-SEP-1997.
PF 10-MAR-1996; 163287.
PA (MEI) : MEIJI MILK PROD CO LTD.
PA (DAI) : DAICEL CHEM IND LTD.
DR WPI: 97-470495/43.
PI Peptide immunotherapeutic agent to treat allergic diseases.
PT Contains multi-epitope peptide containing T cell epitope regions.
PS Claim 1; Page 32; 58pp; Japanese.
CC The present sequence represents a multi-epitope peptide which is used as
CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2
CC or more different allergens (preferably linked via arginine or lysine
CC dimers), where the T cell epitope regions have a positivity index
CC greater than 100 as measured in a patient group responding to the
CC allergen; have at least 70% reactivity with lymphocytes from patients
CC responding to the allergen; and are not reactive with immunoglobulin E
CC (IgE) antibodies from patients responsive to the allergen. The agent can
CC be used to prevent and treat a wide variety of allergic diseases, e.g. by
CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
SQ Sequence 31 AA:

Query Match 93.0% Score 53; DB 1; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.87e-01;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DE 24 IKRVSNVI 31
      |||||
QY 2 IKRVSNVI 9

RESULT 8
ID W27373 standard; peptide: 31 AA.
AC W27373:
DE 24-MAR-1996 (first entry)
DE Multi-epitope peptide used as immunotherapeutic agent #4.
KW Multi-epitope peptide; immunotherapeutic agent; allergic disease.
KW T-cell epitope region; allergen; lymphocyte; immunoglobulin E.
OS Synthetic.
OS Crypthomeria japonica.
PN 21002770-A.
PD 12-SEP-1997.
PF 10-MAR-1996; 163287.
PA (MEI) : MEIJI MILK PROD CO LTD.
PA (DAI) : DAICEL CHEM IND LTD.
DR WPI: 97-470495/43.
PI Peptide immunotherapeutic agent to treat allergic diseases.
PT Contains multi-epitope peptide containing T cell epitope regions.
PS Claim 1; Page 32; 58pp; Japanese.
CC The present sequence represents a multi-epitope peptide which is used as
CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2
CC or more different allergens (preferably linked via arginine or lysine
CC dimers), where the T cell epitope regions have a positivity index
CC greater than 100 as measured in a patient group responding to the
CC allergen; have at least 70% reactivity with lymphocytes from patients
CC responding to the allergen; and are not reactive with immunoglobulin E
CC (IgE) antibodies from patients responsive to the allergen. The agent can
CC be used to prevent and treat a wide variety of allergic diseases, e.g. by
CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
SQ Sequence 31 AA:

Query Match 93.0% Score 53; DB 1; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.87e-01;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DE 24 IKRVSNVI 31
      |||||
QY 2 IKRVSNVI 9

RESULT 9
ID W27369 standard; peptide: 80 AA.
AC W27369:
DE 24-MAR-1996 (first entry)
DE Multi-epitope peptide used as immunotherapeutic agent #4.
KW Multi-epitope peptide; immunotherapeutic agent; allergic disease.
KW T-cell epitope region; allergen; lymphocyte; immunoglobulin E.
OS Synthetic.
OS Crypthomeria japonica.
PN 21002770-A.
PD 12-SEP-1997.
PF 10-MAR-1996; 163287.
PA (MEI) : MEIJI MILK PROD CO LTD.
PA (DAI) : DAICEL CHEM IND LTD.
DR WPI: 97-470495/43.
PI Peptide immunotherapeutic agent to treat allergic diseases.
PT Contains multi-epitope peptide containing T cell epitope regions.
PS Claim 6; Page 31; 58pp; Japanese.
CC The present sequence represents a multi-epitope peptide which is used as
CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2
CC or more different allergens (preferably linked via arginine or lysine
CC dimers), where the T cell epitope regions have a positivity index
CC greater than 100 as measured in a patient group responding to the
CC allergen; have at least 70% reactivity with lymphocytes from patients
CC responding to the allergen; and are not reactive with immunoglobulin E
CC (IgE) antibodies from patients responsive to the allergen. The agent can
CC be used to prevent and treat a wide variety of allergic diseases, e.g. by
CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
SQ Sequence 80 AA:

Query Match 93.0% Score 53; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 1.87e-01;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DE 18 IKRVSNVI 25

```

```

DI 24-MAR-1996 (first entry)
DE Multi-epitope peptide used as immunotherapeutic agent #5.
KW Multi-epitope peptide; immunotherapeutic agent; allergic disease.
KW T-cell epitope region; allergen; lymphocyte; immunoglobulin E.
OS Synthetic.
OS Crypthomeria japonica.
PN 21002770-A.
PD 12-SEP-1997.
PF 10-MAR-1996; 163287.
PA (MEI) : MEIJI MILK PROD CO LTD.
PA (DAI) : DAICEL CHEM IND LTD.
DR WPI: 97-470495/43.
PI Peptide immunotherapeutic agent to treat allergic diseases.
PT Contains multi-epitope peptide containing T cell epitope regions.
PS Claim 9; Page 32; 58pp; Japanese.
CC The present sequence represents a multi-epitope peptide which is used as
CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2
CC or more different allergens (preferably linked via arginine or lysine
CC dimers), where the T cell epitope regions have a positivity index
CC greater than 100 as measured in a patient group responding to the
CC allergen; have at least 70% reactivity with lymphocytes from patients
CC responding to the allergen; and are not reactive with immunoglobulin E
CC (IgE) antibodies from patients responsive to the allergen. The agent can
CC be used to prevent and treat a wide variety of allergic diseases, e.g. by
CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
SQ Sequence 31 AA:

Query Match 93.0% Score 53; DB 1; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.87e-01;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DE 24 IKRVSNVI 31
      |||||
QY 2 IKRVSNVI 9

RESULT 9
ID W27369 standard; peptide: 80 AA.
AC W27369:
DE 24-MAR-1996 (first entry)
DE Multi-epitope peptide used as immunotherapeutic agent #5.
KW Multi-epitope peptide; immunotherapeutic agent; allergic disease.
KW T-cell epitope region; allergen; lymphocyte; immunoglobulin E.
OS Synthetic.
OS Crypthomeria japonica.
PN 21002770-A.
PD 12-SEP-1997.
PF 10-MAR-1996; 163287.
PA (MEI) : MEIJI MILK PROD CO LTD.
PA (DAI) : DAICEL CHEM IND LTD.
DR WPI: 97-470495/43.
PI Peptide immunotherapeutic agent to treat allergic diseases.
PT Contains multi-epitope peptide containing T cell epitope regions.
PS Claim 6; Page 31; 58pp; Japanese.
CC The present sequence represents a multi-epitope peptide which is used as
CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2
CC or more different allergens (preferably linked via arginine or lysine
CC dimers), where the T cell epitope regions have a positivity index
CC greater than 100 as measured in a patient group responding to the
CC allergen; have at least 70% reactivity with lymphocytes from patients
CC responding to the allergen; and are not reactive with immunoglobulin E
CC (IgE) antibodies from patients responsive to the allergen. The agent can
CC be used to prevent and treat a wide variety of allergic diseases, e.g. by
CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
SQ Sequence 80 AA:

Query Match 93.0% Score 53; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 1.87e-01;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DE 18 IKRVSNVI 25

```

```

Query Match          93.0%  Score 53:  DB 1: Length 134
Best Local Similarity 100.0%  Pred. No. 1: 87e-01
Matches      8:  Conservative      0:  Mismatches      0:  Indels      0:  Gaps      0:

Db      18 IKRVSNVI 25
QY      2 IKRVSNVI 9

RESULT 12
ID R75388 standard: protein; 353 AA
AC R75388
DE Japanese cedar pollen allergen Cryj I.
KW Japanese cedar; pollen allergen; Cryj I; T-cell epitope; peptides
KW prevention; treatment; cryptomeria pollinosis.
OS Cryptomeria japonica.
FH Key
FT peptide 61..75 Location/Qualifiers
FT peptide /note="T-cell epitope peptide"
FT peptide 91..105
FT peptide /note="T-cell epitope peptide"
FT peptide 105..120
FT peptide /note="T-cell epitope peptide"
FT peptide 146..160
FT peptide /note="T-cell epitope peptide"
FT peptide 211..225
FT peptide /note="T-cell epitope peptide"
FT peptide 326..340
FT peptide /note="T-cell epitope peptide"
FT peptide 335..346
FT peptide /note="T-cell epitope peptide"
PN J07118295-A.
PD 09-MAY-1995.
PF 20-OCT-1993: 262626.
PR 20-OCT-1993: JP-262626.
PA (MEIP ) MEIJI MILK PROD CO LTD.
ER WPI: 95-203834/27.
PT New cryptomeria pollen allergen T-cell epitope peptide used for
PT prevention, treatment and investigation of Japanese cedar pollinosis.
PS Disclosure: Figs 1-2: epp; Japanese.
CC R75388 is the Japanese cedar pollen allergen Cryj I, from which the
CC T-cell epitope peptides R89289-R89295 were derived. The peptides
CC can be used for the prevention and treatment of cryptomeria
CC pollinosis, and also for the investigation of pollinosis.
SQ Sequence 353 AA:

Query Match          93.0%  Score 53:  DB 1: Length 134
Best Local Similarity 100.0%  Pred. No. 1: 87e-01
Matches      8:  Conservative      0:  Mismatches      0:  Indels      0:  Gaps      0:

Db      110 IKRVSNVI 117
QY      2 IKRVSNVI 9

RESULT 13
ID R81587 standard: Protein; 353 AA
AC R81587
DE 24-MAY-1996 (first entry)
DE Cedar pollen allergen B.
KW Cedar; pollen; allergen; immunoglobulin E; IgE; T-cell epitope;
KW antibody; pollinosis; therapy; immunotherapy.
OS Cryptomeria japonica.
PN EP-700929-A2.
PD 13-MAR-1996.
PF 08-SEP-1995: 305295.
PR 10-SEP-1994: JP-242137.
PR 14-JUL-1995: JP-200221.
PR 14-JUL-1995: JP-200204.
PA (HAYB ) HAYASHIBARA SEIBUTSU KAGAKU.

Query Match          93.0%  Score 53:  DB 1: Length 105
Best Local Similarity 100.0%  Pred. No. 1: 87e-01
Matches      8:  Conservative      0:  Mismatches      0:  Indels      0:  Gaps      0:

Db      18 IKRVSNVI 25
QY      2 IKRVSNVI 9

RESULT 11
ID W27370 standard: peptide; 134 AA
AC W27370
DE 24-MAR-1998 (first entry)
DE Multi-epitope peptide used as immunotherapeutic agent #2.
KW Multi-epitope peptide; immunotherapeutic agent; allergic disease;
KW T-cell epitope region; allergen; lymphocyte; immunoglobulin E.
OS Synthetic.
PN W09732600-A1.
PD 12-SEP-1997.
PF 10-MAR-1997: J00740.
PR 10-MAR-1996: JP-080702.
PA (MEIP ) MEIJI MILK PROD CO LTD.
ER Dai-iki K. Iwama A. Kino K. Kume A. Sone T.
WPI: 97-470495/43.
PT Peptide immunotherapeutic agent to treat allergic diseases -
PT contains multi-epitope peptide containing T cell epitope regions
PS Claim 6: Page 42: 58pp; Japanese.
CC The present sequence represents a multi-epitope peptide which is used as
CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2
CC or more different allergens (preferably linked via arginine or lysine
CC dimers), where the T cell epitope regions have a positivity index
CC greater than 100 as measured in a patient group responding to the
CC allergen; have at least 70% reactivity with lymphocytes from patients
CC responding to the allergen; and are not reactive with immunoglobulin E
CC (IgE) antibodies from patients responsive to the allergen. The agent can
CC be used to prevent and treat a wide variety of allergic diseases, e.g. by
CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
SQ Sequence 105 AA:

Query Match          93.0%  Score 53:  DB 1: Length 105
Best Local Similarity 100.0%  Pred. No. 1: 87e-01
Matches      8:  Conservative      0:  Mismatches      0:  Indels      0:  Gaps      0:

Db      18 IKRVSNVI 25
QY      2 IKRVSNVI 9

RESULT 11
ID W27370 standard: peptide; 134 AA
AC W27370
DE 24-MAR-1998 (first entry)
DE Multi-epitope peptide used as immunotherapeutic agent #3.
KW Multi-epitope peptide; immunotherapeutic agent; allergic disease;
KW T-cell epitope region; allergen; lymphocyte; immunoglobulin E.
OS Synthetic.
PN W09732600-A1.
PD 12-SEP-1997.
PF 10-MAR-1997: J00740.
PR 10-MAR-1996: JP-080702.
PA (MEIP ) MEIJI MILK PROD CO LTD.
ER Dai-iki K. Iwama A. Kino K. Kume A. Sone T.
WPI: 97-470495/43.
PT Peptide immunotherapeutic agent to treat allergic diseases -
PT contains multi-epitope peptide containing T cell epitope regions
PS Claim 6: Page 42: 58pp; Japanese.
CC The present sequence represents a multi-epitope peptide which is used as
CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2
CC or more different allergens (preferably linked via arginine or lysine
CC dimers), where the T cell epitope regions have a positivity index
CC greater than 100 as measured in a patient group responding to the
CC allergen; have at least 70% reactivity with lymphocytes from patients
CC responding to the allergen; and are not reactive with immunoglobulin E
CC (IgE) antibodies from patients responsive to the allergen. The agent can
CC be used to prevent and treat a wide variety of allergic diseases, e.g. by
CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
SQ Sequence 105 AA:

```


PI Hino K., Saito S., Taniguchi Y.
 DR WPI: 96-14076/45.
 PT New peptides derived from cedar pollen allergens - activate
 PT Allergen-specific T-cells, but not allergen-specific IgE antibodies,
 PT used for treating cedar pollinosis.
 PS Claim 5: Page 31-32; 36pp; English.
 CC Synthetic peptides based on portions of cedar pollen allergens A
 CC (R8156) and B (R8158) were tested for their ability to activate
 CC cedar allergen-specific T-cells, but not allergen-specific IgE
 CC antibodies. 5 peptides (R8158-R81585) were identified as T-cell
 CC epitopes. These peptides, plus subsequent peptides (R8158-96) can
 CC for T-cell reactivation, and homologous peptides (R8158-96) can
 CC be used as immunotherapeutic agents to treat or prevent cedar
 CC pollinosis, avoiding side-effects such as anaphylaxis.
 SQ Sequence 374 AA.

Query Match 93.0% Score 53 DB 17 Length 374

Best Local Similarity 100.0% Pred. No. 1.87e+01

Matches 8: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

DB 131 IKRVSNI: 132

QY 2 IKRVSNI: 9

RESULT 14
 IC R60164 standard: Protein: 374 AA.

AC R60164

DT 24-MAR-1995 (first entry)

DE Japanese cedar pollen antigen CryjI.

KW Japanese cedar: pollen antigen: allergen: CryjI: sig: pollinosis.

CS Cryptomeria japonica.

FR Key Location/Qualifiers

FI peptide 1..21

FI protein 22..374

FI /label= signal_peptide

FI /label= mature_CryjI

PN JCM97768-A.

PD 19-JUL-1994

PF 07-JAN-1993 G01116.

PR 07-JAN-1993 JP-C0116.

PA (MEJ) MFI: SEIKA KAISHA.

DR WPI: 94-263680/23.

DN N-PSDB: Q71501.

PI Sugi (Japanese cedar) pollen antigen CryjI - is useful for

PI diagnosis, treatment and prevention of sugi pollinosis

PS Claim 4: Page 5-7; 9pp; Japanese.

CC the coding sequence for the Japanese cedar ("sugi") pollen allergen

CC CryjI was isolated from a cDNA library prepared from POLYA mRNA. All

CC or part of the CryjI protein can be used for diagnosis treatment

CC and prevention of sugi pollinosis.

SQ Sequence 374 AA.

Query Match

Best Local Similarity 100.0% Pred. No. 1.87e+01

Matches 8: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

DB 131 IKRVSNI: 138

QY 2 IKRVSNI: 9

RESULT 15

IC R31937 standard: Protein: 374 AA.

AC R31937

DT 03-JUN-1993 (first entry)

DE Cry j I.

KW Japanese cedar pollen: allergen: antigen: allergy: B cell: T cell.

CS Cryptomeria japonica.

FR Key Location/Qualifiers

FI peptide 1..21

FI /note= "signal peptide"

FI protein 22..374

PT /note= "mature Cry j I"

DR W09101213-A.

PD 21-JAN-1993

PF 10-JUL-1992; U05661.

PR 12-JUL-1991; US-729134.

PA 15-JUL-1991; US-730452.

PS (IMMU) IMMUNOLOGIC PHARM CORP.

PI Bond JF, Griffith LJ, Pollock J.

DR WPI: 93-04544/05.

DN N-PSDB: Q35304.

PI Nucleic acid sequence encoding Cryptomeria japonica allergen

PI for the diagnosis treatment and prevention of allergic reactions

PI to Japanese cedar pollen

PS Claim 1: Page 42; 69pp; English

CC Fresh pollen and staminate cone samples were collected from a stand

CC Cryptomeria japonica (Japanese cedar) tree. RNA was prep and sub

CC to synthesize cDNA. The cDNA was subjected to successive rounds of

CC PCR to yield a full length Cry j I clone. Cry j I or an antigenic

CC fragment of it may be used for detecting, treating and preventing an

CC allergic response to Japanese cedar pollen allergen. It is capable of

CC modifying both the B and T cell response to Cry j I and I cell response

CC to a Cry j I antigen.

SQ Sequence 374 AA.

Query Match

Best Local Similarity 93.0% Score 53 DB 17 Length 374

Matches 8: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

DB 131 IKRVSNI: 138

QY 2 IKRVSNI: 9

Search completed: Mon Jun 19 16:19:22 2000

Job time: 11 secs.

Release 3.1A John F. Collins, BioComputing Research Unit,
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

(TM)

WATERM

Wparch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Jun 20 13:37:46 2000; MasPar time 2.17 Seconds

Tabular output not generated. 59,945 Million cell updates/sec

Title: >US-09-142-524A-14
Description: (1-9) from US05142524A.pep
Perfect Score: 57
Sequence: 1 SIKRVSNV: 9

Scoring table: FAM 150
Gap 15

Searched: 145341 seqs, 14437480 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-issued

1:5A_COMB 2:5B_COMB 3:6_COMB 4:PCT_COMB 5:Backfiles1

Statistics: Mean 14.393; Variance 44.43; scale 0.324

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | LR | ID | Description | Pred. No. |
|------------|-------|---------------|--------|----|------------|------------------------|-----------|
| 1 | 44 | 77.2 | 564 | 2 | US-08-565- | Sequence 1, Applicatio | 1.39e+02 |
| 2 | 44 | 77.2 | 610 | 2 | US-08-565- | Sequence 6, Applicatio | 1.39e+02 |
| 3 | 42 | 73.7 | 255 | 1 | US-08-459- | Sequence 4, Applicatio | 2.27e+02 |
| 4 | 42 | 73.7 | 255 | 1 | US-08-459- | Sequence 4, Applicatio | 2.27e+02 |
| 5 | 41 | 71.9 | 357 | 1 | US-08-356- | Sequence 2, Applicatio | 2.90e+02 |
| 6 | 41 | 71.9 | 357 | 1 | US-08-356- | Sequence 4, Applicatio | 2.90e+02 |
| 7 | 41 | 71.9 | 401 | 2 | US-08-576- | Sequence 45, Applicati | 2.90e+02 |
| 8 | 41 | 71.9 | 445 | 2 | US-08-679- | Sequence 3, Applicati | 2.90e+02 |
| 9 | 41 | 71.9 | 2544 | 2 | US-08-679- | Sequence 32, Applicati | 2.90e+02 |
| 10 | 40 | 70.2 | 332 | 2 | US-08-672- | Sequence 8, Applicatio | 3.69e+02 |
| 11 | 40 | 70.2 | 332 | 2 | US-08-780- | Sequence 2, Applicatio | 3.69e+02 |
| 12 | 40 | 70.2 | 332 | 2 | US-08-842- | Sequence 8, Applicatio | 3.69e+02 |
| 13 | 40 | 70.2 | 332 | 2 | US-08-842- | Sequence 8, Applicatio | 3.69e+02 |
| 14 | 40 | 70.2 | 332 | 1 | US-08-671- | Sequence 8, Applicatio | 3.69e+02 |
| 15 | 40 | 70.2 | 885 | 4 | PCT-US95-0 | Sequence 2, Applicatio | 3.69e+02 |
| 16 | 40 | 70.2 | 885 | 4 | PCT-US95-0 | Sequence 2, Applicatio | 3.69e+02 |
| 17 | 40 | 70.2 | 885 | 2 | PCT-US95-0 | Sequence 2, Applicatio | 3.69e+02 |
| 18 | 40 | 70.2 | 909 | 4 | PCT-US95-0 | Sequence 142, Applicat | 3.69e+02 |
| 19 | 40 | 70.2 | 909 | 2 | US-08-310- | Sequence 142, Applicat | 3.69e+02 |
| 20 | 39 | 68.4 | 33 | 4 | PCT-US95-0 | Sequence 97, Applicati | 4.68e+02 |
| 21 | 39 | 68.4 | 33 | 4 | PCT-US95-0 | Sequence 97, Applicati | 4.68e+02 |
| 22 | 39 | 68.4 | 33 | 2 | US-08-310- | Sequence 97, Applicati | 4.68e+02 |
| 23 | 39 | 68.4 | 227 | 1 | US-08-745- | Sequence 6, Applicatio | 4.68e+02 |

| | | | | | | | |
|----|----|------|------|---|------------|------------------------|----------|
| 24 | 39 | 68.4 | 227 | 1 | US-08-597- | Sequence 6, Applicatio | 4.68e+02 |
| 25 | 39 | 68.4 | 317 | 2 | US-08-518- | Sequence 2, Applicatio | 4.68e+02 |
| 26 | 39 | 68.4 | 317 | 2 | US-08-518- | Sequence 2, Applicatio | 4.68e+02 |
| 27 | 39 | 68.4 | 317 | 3 | US-08-518- | Sequence 2, Applicatio | 4.68e+02 |
| 28 | 39 | 68.4 | 346 | 1 | US-07-661- | Sequence 4, Applicatio | 4.68e+02 |
| 29 | 39 | 68.4 | 503 | 4 | PCT-US95-0 | Sequence 8, Applicatio | 4.68e+02 |
| 30 | 39 | 68.4 | 503 | 1 | US-08-307- | Sequence 6, Applicatio | 4.68e+02 |
| 31 | 39 | 68.4 | 503 | 1 | US-08-474- | Sequence 8, Applicatio | 4.68e+02 |
| 32 | 39 | 68.4 | 503 | 1 | US-08-245- | Sequence 8, Applicatio | 4.68e+02 |
| 33 | 39 | 68.4 | 1019 | 2 | US-08-222- | Sequence 26, Applicati | 4.68e+02 |
| 34 | 39 | 68.4 | 1019 | 1 | US-08-271- | Sequence 7, Applicatio | 4.68e+02 |
| 35 | 38 | 66.7 | 12 | 1 | US-07-778- | Sequence 34, Applicati | 5.94e+02 |
| 36 | 38 | 66.7 | 12 | 1 | US-07-778- | Sequence 34, Applicati | 5.94e+02 |
| 37 | 38 | 66.7 | 192 | 1 | US-08-085- | Sequence 80, Applicati | 5.94e+02 |
| 38 | 38 | 66.7 | 192 | 4 | US-08-085- | Sequence 80, Applicati | 5.94e+02 |
| 39 | 38 | 66.7 | 585 | 4 | US-08-151- | Sequence 3, Applicatio | 5.94e+02 |
| 40 | 38 | 66.7 | 641 | 1 | US-07-718- | Sequence 3, Applicatio | 5.94e+02 |
| 41 | 38 | 66.7 | 641 | 1 | US-07-718- | Sequence 3, Applicatio | 5.94e+02 |
| 42 | 38 | 66.7 | 681 | 2 | US-08-555- | Sequence 4, Applicatio | 5.94e+02 |
| 43 | 38 | 66.7 | 681 | 4 | PCT-US96-0 | Sequence 4, Applicatio | 5.94e+02 |
| 44 | 38 | 66.7 | 803 | 5 | US-08-745- | Sequence 2, Applicatio | 5.94e+02 |
| 45 | 38 | 66.7 | 807 | 4 | PCT-US96-0 | Sequence 2, Applicatio | 5.94e+02 |

ALIGNMENTS

RESULT 1
ID US-08-565-337A-1 STANDARD PRT: 564 AA
XX
AC xxxxxx
XX

Sequence 1, Application US/-8865337A

Sequence 2, Application US/8865337A

Patent No. 5972649

GENERAL INFORMATION:

APPLICANT: Ac-Young, Janice

APPLICANT: Covitz, Peter

APPLICANT: Tang, Y. Tcm

APPLICANT: Murphy, Lynn

TITLE OF INVENTION: HUMAN MULTIPLE ENDONUCLEASE PLASMA TYPE

TITLE OF INVENTION: PROTEIN

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals Inc.

STREET: 3174 Porter Ave

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/865,337A

FILING DATE: Herewith

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy

REGISTRATION NUMBER: 749

REFERENCE/DOCKET NUMBER: PF-0305 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 1:

| CC | SEQUENCE | 610 AA: 67456 MW: 1878916 CN: |
|--------|--|--|
| CC | CLONE: 1945387 | |
| CC | SEQUENCE | 610 AA: 67456 MW: 1878916 CN: |
| CC | Query Match | 77.2% Score 44: DB 2: Length 610 |
| CC | Best Local Similarity | 62.5% Pred. No. 1390-02 |
| CC | Matches | 5: Conservative 0: Mismatches 0: Indels 0: Gaps 0: |
| DB | 118 VKXSVLVI 125 | |
| QY | 2 IKRVSNI 9 | |
| RESULT | 3 | |
| ID | US-08-459-264-4 | STANDARD: PRI: 255 AA: |
| XX | xxxxxx | |
| AC | | |
| XX | | |
| DT | | |
| XX | | |
| DE | Sequence 4, Application US/0445-264 | |
| XX | Sequence 4, Application US/ 459264 | |
| CC | Patent No. 5705340 | |
| CC | GENERAL INFORMATION: | |
| CC | APPLICANT: RASMUSSEN, B. L. A. | |
| CC | APPLICANT: TALLY, FRANK S. P. | |
| CC | APPLICANT: GLUZMAN, YAKO | |
| CC | TITLE OF INVENTION: DNA | |
| CC | TITLE OF INVENTION: OF CLASS B BETA-LACTAMASE ENZYMES FROM BACTERIES | |
| CC | TITLE OF INVENTION: FRAC 115 | |
| CC | NUMBER OF SEQUENCES: 9 | |
| CC | CORRESPONDENCE ADDRESS: | |
| CC | ADDRESSER: Darby & Darby Inc. | |
| CC | STREET: 805 Third Avenue | |
| CC | CITY: New York | |
| CC | STATE: NY | |
| CC | COUNTRY: US | |
| CC | ZIP: 10022 | |
| CC | COMPUTER READABLE FORM: | |
| CC | MEDIUM TYPE: Floppy disk | |
| CC | COMPUTER: IBM PC compatible | |
| CC | OPERATING SYSTEM: PC-DOS/MS-DOS | |
| CC | SOFTWARE: Patent in Release #10, Version #1.0 | |
| CC | CURRENT APPLICATION DATA: | |
| CC | APPLICATION NUMBER: US/8/459,264 | |
| CC | FILING DATE: 02-JUN-1991 | |
| CC | CLASSIFICATION: 435 | |
| CC | ATTORNEY/AGENT INFORMATION: | |
| CC | NAME: PATENT INVENTOR | |
| CC | REGISTRATION NUMBER: 44/19-1-121 | |
| CC | TELECOMMUNICATION INFORMATION: | |
| CC | TELEPHONE: 212-527-7700 | |
| CC | TELEFAX: 212-753-6237 | |
| CC | TELEX: 236687 | |
| CC | INFORMATION FOR SEQ ID NO: 1: | |
| CC | SEQUENCE CHARACTERISTICS: | |
| CC | LENGTH: 255 amino acids | |
| CC | TYPE: amino acid | |
| CC | STRANDEDNESS: single | |
| CC | TOPOLOGY: linear | |
| CC | MOLECULE TYPE: peptide | |
| CC | ORIGINAL SOURCE: | |
| CC | ORGANISM: Bacillus cereus | |
| CC | IMMEDIATE SOURCE: | |
| CC | CLONE: B-LACTAMASE | |
| CC | SEQUENCE 255 AA: 27847 MW: 349604 CN: | |
| QY | Sequence Match | 73.7% Score 42: DB 1: Length 255: |
| CC | Best Local Similarity | 71.4% Pred. No. 2,278-02: |
| CC | Matches | 5: Conservative 2: Mismatches 0: Indels 0: Gaps 0: |
| DB | 105 KSVTVNI 111 | |

```

QY      3 KRVSNI 9
      IIIII
RESULT  4
ID      US-08-459-263-4      STANDARD:      PRT:      255 AA.
XX      XXXXXX
XX
XX
DE      Sequence 4, Application US/08459263
XX      Sequence 4, Application US/08459263
CC      Patent No. 5792642
CC      GENERAL INFORMATION:
CC      APPLICANT: RASMUSSEN, Beth A
CC      APPLICANT: TALLY, Francis P
CC      APPLICANT: GLUZMAN, Yakov
CC      TITLE OF INVENTION: DNA SEQUENCES AND AMINO ACID SEQUENCES
CC      TITLE OF INVENTION: OF CLASS B BETA-LACTAMASE ENZYMES FROM BACTERIOIDES
CC      TITLE OF INVENTION: FRAGILIS
CC      NUMBER OF SEQUENCES: 9
CC      CORRESPONDENCE ADDRESS:
CC      ADDRESSEE: Darby & Darby PC
CC      STREET: 805 Third Avenue
CC      CITY: New York
CC      STATE: NY
CC      COUNTRY: US
CC      ZIP: 10022
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: Patent In Release #1.30, Version #1.30
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: US/08/459,263
CC      FILING DATE: 02-JUN-1995
CC      CLASSIFICATION: 435
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: Robinson, Joseph R
CC      REGISTRATION NUMBER: 33,448
CC      REFERENCE/DOCKET NUMBER: 0646/1B026-US2
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: 212-527-7700
CC      TELEFAX: 212-753-6237
CC      TELEX: 234687
CC      INFORMATION FOR SEQ ID NO: 4:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 255 amino acids
CC      TYPE: amino acid
CC      STRANDEDNESS: single
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: peptide
CC      ORIGINAL SOURCE:
CC      ORGANISM: Bacillus cereus
CC      IMMEDIATE SOURCE:
CC      CLONE: B-LACTAMASE
CC      SEQUENCE 255 AA: 27647 MW: 349604 CN:

Query Match 73.7% Score 42: DB 1: Length 255:
Best Local Similarity 71.4% Pred. NO. 2,278-02:
Matches 5: Conservative 2: Mismatches 0: Indels 0: Gaps 0:

DB 105 KRVTDV: IIII
QY 3 KRVSNI 9
      IIIII
RESULT  5
ID      US-08-356-405-2      STANDARD:      PRT:      357 AA.
XX      XXXXXX
XX
XX
DE      Sequence 4, Application US/08031538
XX      Sequence 4, Application US/08031538
CC      Patent No. 5968817
CC
CC
DE      Sequence 2, Application US/08356405
XX      Sequence 2, Application US/08356405
CC      Patent No. 5807691
CC      GENERAL INFORMATION:
CC      APPLICANT: Amalaky, No. 5807691rdine
CC      APPLICANT: Boschert, Ursula
CC      APPLICANT: Hen, Rene
CC      APPLICANT: Plassat, Jean-Luc
CC      TITLE OF INVENTION: Polypeptides Having Serotonin Receptor
CC      TITLE OF INVENTION: Activity (SHT5A), Nucleic Acids Coding for These
CC      TITLE OF INVENTION: Polypeptides and Uses Thereof:
CC      NUMBER OF SEQUENCES: 13
CC      CORRESPONDENCE ADDRESS:
CC      ADDRESSEE: Rhone-Poulenc Rorer Inc.
CC      STREET: 500 Arcola Road, 3043
CC      CITY: Collegeville
CC      STATE: PA
CC      COUNTRY: USA
CC      ZIP: 19426-0107
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: Patent In Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: US/08/356,405
CC      FILING DATE:
CC      CLASSIFICATION: 435
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: WO PCT/EP91/00650
CC      FILING DATE: 29-JUN-1992
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: FR 92/06091
CC      FILING DATE: 01-JUL-92
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: Smith, Julie K
CC      REGISTRATION NUMBER: 29,613
CC      REFERENCE/DOCKET NUMBER: EX92004-US
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: (610)454-3849
CC      TELEFAX: (610)454-3808
CC      INFORMATION FOR SEQ ID NO: 2:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 357 amino acids
CC      TYPE: amino acid
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: protein
CC      SEQUENCE 357 AA: 43804 MW: 741957 CN:

Query Match 71.9% Score 41: DB 1: Length 357:
Best Local Similarity 85.7% Pred. NO. 2,908-02:
Matches 6: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

DB 155 KRVSVM 151
      IIIII
QY 3 KRVSNI 9
      IIIII
RESULT  6
ID      US-08-031-538-4      STANDARD:      PRT:      357 AA.
XX      XXXXXX
XX
XX
DE      Sequence 4, Application US/08031538
XX      Sequence 4, Application US/08031538
CC      Patent No. 5968817
CC
CC

```

CC GENERAL INFORMATION:
 CC APPLICANT: Sutcliffe, J. Gregor
 CC APPLICANT: Brander, Mark G
 CC APPLICANT: Lovelace, Timothy W
 CC TITLE OF INVENTION: HUMAN SEROTIN RECEPTORS, DNA ENCODING
 CC TITLE OF INVENTION: THE RECEPTORS, AND USES THEREOF
 CC NUMBER OF SEQUENCES: 17
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESS: The Scripps Research Institute, Office of
 CC ADDRESS: Patent Counsel
 CC STREET: 10555 NO. 546217th Torrey Pines Road, LPO B
 CC CITY: La Jolla
 CC STATE: CA
 CC COUNTRY: USA
 CC ZIP: 92037
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patent Release #1.0, Version #1.45
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/041,538
 CC FILING DATE: 1998/03/15
 CC CLASSIFICATION: B01
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Filling, Thomas
 CC REGISTRATION NUMBER: 34,163
 CC REFERENCE/DOCKET NUMBER: TSP5099P
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: 619-554-2937
 CC TELEFAX: 619-554-6412
 CC INFORMATION FOR SEQ ID NO: 4:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 350 amino acids
 CC TYPE: AMINO ACIDS
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC SEQUENCE 4: AA: 43672 MW: 736.46 CN:
 CC
 CC Query Match 71.9% Score 411 BB 2 Length 457
 CC Best Local Similarity 85.7% Pred. No. 2,900-021
 CC Matches 60 Conservative 11 Mismatches 60 Indels 00 Gaps 00
 CC
 CC DB 150 KRVSNN 151
 CC 11111
 CC 3 KRVSNN
 CC
 CC RESULT 7
 CC ID US-09-576-266-10 Score 411 BB 2 Length 457
 CC AC xxxxxx
 CC DE
 CC Sequence 45: Application US/08/576266A
 CC
 CC Sequence 45: Application US/08/576266A
 CC Patent No. 59-8194
 CC GENERAL INFORMATION:
 CC APPLICANT: Sutcliffe, J.G.
 CC APPLICANT: Katz, L
 CC APPLICANT: Donadio, S.
 CC APPLICANT: Staver, N.J.
 CC TITLE OF INVENTION: POLYKETIDE-ASSOCIATED SUGAR
 CC TITLE OF INVENTION: BIOSYNTHESIS GENES
 CC NUMBER OF SEQUENCES: 50
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESS: Abbott Laboratories
 CC STREET: 100 Abbott Park Road
 CC CITY: Abbott Park
 CC STATE: Illinois
 CC COUNTRY: USA

CC ZIP: 60064-3500
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Diskette
 CC COMPUTER: IBM Compatible
 CC OPERATING SYSTEM: DOS
 CC SOFTWARE: FastSeq Version 4.0
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/576,426A
 CC FILING DATE: 21-DEC-1998
 CC CLASSIFICATION: 435
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER:
 CC FILING DATE:
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Diagne Casito
 CC REGISTRATION NUMBER: P-40,944
 CC REFERENCE/DOCKET NUMBER: US/08/05,001
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (847) 938-3137
 CC TELEFAX: (847) 938-2623
 CC TELEX:
 CC INFORMATION FOR SEQ ID NO: 45:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 401 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS: single
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: No. 5998194P
 CC SEQUENCE 401: AA: 43347 MW: 755240 CN:
 CC
 CC Query Match 71.9% Score 411 BB 2 Length 457
 CC Best Local Similarity 50.0% Pred. No. 2,900-021
 CC Matches 40 Conservative 30 Mismatches 10 Indels 00 Gaps 00
 CC
 CC DB 372 IRRVCDII 379
 CC 11111
 CC 2 IRRVSNV: 9
 CC
 CC RESULT 6
 CC ID US-08-679-025A-3 STANDARD 1970 445 AA
 CC AC xxxxxx
 CC DE
 CC Sequence 4: Application US/08/679025A
 CC
 CC Sequence 4: Application US/08/679025A
 CC Patent No. 59-8194
 CC GENERAL INFORMATION:
 CC APPLICANT: Sutcliffe, J.G.
 CC APPLICANT: Lovelace, Timothy W
 CC APPLICANT: Brander, Mark G
 CC TITLE OF INVENTION: AUXILIARY GENES AND PROTEINS OF
 CC TITLE OF INVENTION: MEMBRANOLIN RESISTANT BACTERIA AND ANALOGS THEREOF
 CC NUMBER OF SEQUENCES: 17
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESS: David A. Jackson, Esq.
 CC STREET: 411 Hackensack Ave., Commercial Plaza, 4th
 CC CITY: Hackensack
 CC STATE: New Jersey
 CC COUNTRY: USA
 CC ZIP: 07601
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patent Release #1.0, Version #1.30
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/679,535A
 CC FILING DATE: 10-JUL-1998
 CC CLASSIFICATION: 435

CC ATTORNEY/AGENT INFORMATION:
CC NAME: Jackson Esq., David A.
CC REGISTRATION NUMBER: 26,742
CC REFERENCE/DOCKET NUMBER: 600-1-141
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 201-487-5800
CC TELEFAX: 201-343-1584
CC INFORMATION FOR SEQ ID NO: 3:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 445 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC HYPOTHETICAL: NO
CC SEQUENCE 445 AA: 49222 MW: 983134 CN:

Query Match 71.9% Score 41: DB 2: Length 445;
Best Local Similarity 50.0% Pred. No. 2.90e+02;
Matches 5: Conservative 2: Mismatches 2: Indels 0: Gaps 0:

DB 149 SAKRIDDI 156
QY 1 SIKRVSNVI 9

RESULT 9
ID US-08-576-626A-32 STANDARD: PRT: 2544 AA:

XX xxxxxx

Sequence 32, Application: US/08576626A

CC Patent No. 5998194
CC GENERAL INFORMATION:
CC APPLICANT: Summers, R.O.
CC APPLICANT: Katz, L.
CC APPLICANT: Donadio, S.
CC APPLICANT: Slaver, M.J.
CC TITLE OF INVENTION: POLYKETIDE-ASSOCIATED SUGAR
CC TITLE OF INVENTION: BIOSYNTHESIS GENES
CC NUMBER OF SEQUENCES: 60
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Abbott Laboratories
CC STREET: 100 Abbott Park Road
CC CITY: Abbott Park
CC STATE: Illinois
CC COUNTRY: USA
CC ZIP: 60064-3500
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette
CC COMPUTER: IBM Compatible
CC OPERATING SYSTEM: DOS
CC SOFTWARE: FastSeq Version 2.0
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/576.626A
CC FILING DATE: 21-DEC-1995
CC CLASSIFICATION: 435
CC PRIOR APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Dianne Casuto
CC REGISTRATION NUMBER: P-40,943
CC REFERENCE/DOCKET NUMBER: 5857.US.01
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (847) 938-3137
CC TELEFAX: (847) 938-2623
CC TELEX:
CC INFORMATION FOR SEQ ID NO: 32:

CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 2544 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: NO: 5998194e
SQ SEQUENCE 2544 AA: 279108 MW: 3:031572 CN:

Query Match 71.9% Score 41: DB 2: Length 2544;
Best Local Similarity 50.0% Pred. No. 2.97e+02;
Matches 4: Conservative 3: Mismatches 1: Indels 0: Gaps 0:

DB 1833 IRRVCDII 1840
QY 2 IKRVSNVI 9

RESULT 10
ID US-08-672-109B-8 STANDARD: PRT: 332 AA:

XX xxxxxx

Sequence 8, Application: US/08 72:09B

CC Sequence 8, Application US/08672109B
CC Patent No. 5710265
CC GENERAL INFORMATION:
CC APPLICANT: Yamada, Tad. Ka
CC APPLICANT: Gantz, Ira
CC TITLE OF INVENTION: Gene Encoding Melanocortin Receptors
CC NUMBER OF SEQUENCES: 23
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
CC STREET: P.O. Box 892
CC CITY: Bloomfield Hills
CC STATE: MI
CC COUNTRY: US
CC ZIP: 48303
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/672.109B
CC FILING DATE: June 27, 1994
CC CLASSIFICATION: 536
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Smith, DeAnn F.
CC REGISTRATION NUMBER: 36693
CC REFERENCE/DOCKET NUMBER: 2115-000893DVC
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (810)641-1600
CC TELEFAX: (810)641-0270
CC INFORMATION FOR SEQ ID NO: 8:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 332 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
SQ SEQUENCE 332 AA: 36956 MW: 598933 CN:

Query Match 70.2% Score 40: DB 1: Length 332;
Best Local Similarity 44.4% Pred. No. 3.69e+02;
Matches 4: Conservative 4: Mismatches 1: Indels 0: Gaps 0:

DB 162 TVKRVGIII 170
QY 1 SIKRVSNVI 9

CC FILING DATE:
 CC CLASSIFICATION: 536
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Smith, Deann F.
 CC REGISTRATION NUMBER: 36583
 CC REFERENCE/DOCKET NUMBER: 2115-000853DVB
 CC TELEPHONE: (810)641-1600
 CC TELEFAX: (810)641-0270
 CC INFORMATION FOR SEQ ID NO: 8:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 332 amino acids
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC SEQUENCE 332 AA: 36956 MW: 598933 CN:
 Query Match 70.2% Score 40: DB 2: Length 332:
 Best Local Similarity 44.4% Pred. No. 3.69e-02:
 Matches 4: Conservative 4: Mismatches 1: Indels 0:
 Db 162 TVKRVGIII 170
 QY 1 SIKRVSNVI 9
 RESULT 14
 ID US-08-671-5255-8 STANDARD: PRT: 332 AA:
 XX XXXXXX
 DE Sequence 8, Application US/08671525B
 CC Patent No. 5753220
 CC GENERAL INFORMATION:
 CC APPLICANT: Yamada, Tadataka
 CC APPLICANT: Gantz, Ira
 CC TITLE OF INVENTION: Genes Encoding Melanocortin Receptors
 CC NUMBER OF SEQUENCES: 23
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Hatness, Dickey & Pierce, P.L.C.
 CC STREET: P.O. Box 828
 CC CITY: Bloomfield Hills
 CC STATE: MI
 CC COUNTRY: US
 CC ZIP: 48303
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patent Release #1.0, Version #1.25
 CC CURRENT APPLICATION NUMBER: US/08/671-525B
 CC FILING DATE: June 27, 1996
 CC CLASSIFICATION: 435
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Smith, Deann F.
 CC REGISTRATION NUMBER: 36583
 CC REFERENCE/DOCKET NUMBER: 2115-000853DVB
 CC TELEPHONE: (810)641-1600
 CC TELEFAX: (810)641-0270
 CC INFORMATION FOR SEQ ID NO: 8:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 332 amino acids
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC SEQUENCE 332 AA: 36956 MW: 598933 CN:

CC Query Match 70.2% Score 40: DB 1: Length 332:
 CC Best Local Similarity 44.4% Pred. No. 3.69e-02:
 CC Matches 4: Conservative 4: Mismatches 1: Indels 0:
 Db 162 TVKRVGIII 170
 QY 1 SIKRVSNVI 9
 RESULT 15
 ID PCI-US95-04589-2 STANDARD: PRT: 885 AA:
 XX XXXXXX
 DE Sequence 2, Application PC/US9504589
 CC Sequence 2, Application PC/US9504589
 CC GENERAL INFORMATION:
 CC APPLICANT: Ausubel, Frederick M.
 CC APPLICANT: Staskiewicz, Brian C.
 CC APPLICANT: Brent, Andrew F.
 CC APPLICANT: Dahbeck, Douglas
 CC APPLICANT: Katagiri, Fumaki
 CC APPLICANT: Kunkel, Barbara N.
 CC APPLICANT: Mindrinos, Michael N.
 CC APPLICANT: Yu, Guo-Liuh
 CC TITLE OF INVENTION: RAS GENE AND USES THEREOF
 CC NUMBER OF SEQUENCES: 20
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Fish & Richardson
 CC STREET: 225 Franklin Street Suite 3100
 CC CITY: Boston
 CC STATE: MA
 CC COUNTRY: USA
 CC ZIP: 02110-2904
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC OPERATING SYSTEM: IBM PC compatible
 CC SOFTWARE: Patent Release #1.0, Version #1.30B
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: PCI/US95/04589
 CC FILING DATE:
 CC CLASSIFICATION:
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: US 08/227,345
 CC FILING DATE: 13-APR-1994
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Clark, Paul J.
 CC REGISTRATION NUMBER: 36112
 CC REFERENCE/DOCKET NUMBER: 08786/210001
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (617) 542-5070
 CC TELEFAX: (617) 542-8906
 CC TELEX: 100254
 CC INFORMATION FOR SEQ ID NO: 2:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 885 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS: not relevant
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC SEQUENCE 885 AA: 102110 MW: 4152599 CN:
 Query Match 70.2% Score 40: DB 4: Length 885:
 Best Local Similarity 44.4% Pred. No. 3.69e-02:
 Matches 4: Conservative 5: Mismatches 0: Indels 0:
 Db 325 SIKRLAEII 333
 QY 1 SIKRVSNVI 9

Search completed: Tue Jun 20 13:37:51 2000
JCL time : 5 secs.

CC APPLICATION NUMBER: 07/938,990
 CC FILING DATE: September 1, 1992
 CC APPLICATION NUMBER: PCT/US93/00139
 CC FILING DATE: January 15, 1993
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Darlene A. Vanstone
 CC REGISTRATION NUMBER: 35,729
 CC REFERENCE/DOCKET NUMBER: 025.6 US (IM1-029072)
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (617) 466-6000
 CC TELEFAX: (617) 466-6000
 CC INFORMATION FOR SEQ ID NO: 36:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 20 amino acids
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: polypeptide
 CC FRAGMENT TYPE: internal
 CC SEQUENCE: 20 AA: 2090 MW: 2304 CN:
 Query Match 93.0% Score 53 DB 10 Length 20:
 Best Local Similarity 100.0% Pred. No. 4,33e-017
 Matches 8: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 Db 10 IKRVSNI 17
 III III
 QY 2 IKRVSNI 4
 XXXXX
 RESULT 2
 ID US-09-467-006-36 STANDARD PRI: 20 AA:
 XX XXXXX
 AC XXXXX
 DE Sequence 36, Application US/09467006
 Sequence 36, Application US/09467006
 GENERAL INFORMATION:
 CC APPLICANT: Griffith, Irwin J.
 CC APPLICANT: Pollock, Joanne
 CC APPLICANT: Bond, Julian F.
 CC APPLICANT: Garman, Richard D.
 CC APPLICANT: Kuo, Mei-Chang
 CC APPLICANT: Yeung, Su-mei H.
 CC APPLICANT: Blauer, Andrew
 CC APPLICANT: Exley, Mark A.
 CC APPLICANT: Powers, Steven P.
 CC TITLE OF INVENTION: Allergic Proteins And Peptides From
 CC NUMBER OF SEQUENCES: 46
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
 CC STREET: 610 Lincoln St
 CC CITY: Waltham
 CC STATE: MA
 CC COUNTRY: USA
 CC ZIP: 02154
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC OPERATING SYSTEM: IBM PC compatible
 CC SOFTWARE: Patent Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC FILING DATE: June 6, 1995
 CC CLASSIFICATION: 424
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/350,225
 CC FILING DATE: December 6, 1994
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Jane E. Remillard

CC REGISTRATION NUMBER: 38,872
 CC REFERENCE/DOCKET NUMBER: 025.6 US06 (IM1-028006)
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (617) 227-7400
 CC TELEFAX: (617) 227-5941
 CC INFORMATION FOR SEQ ID NO: 36:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 20 amino acids
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: peptide
 CC FRAGMENT TYPE: internal
 CC SEQUENCE: 20 AA: 2090 MW: 2304 CN:
 Query Match 93.0% Score 54 DB 10 Length 20:
 Best Local Similarity 100.0% Pred. No. 4,33e-017
 Matches 8: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 Db 10 IKRVSNI 17
 III III
 QY 2 IKRVSNI 9
 XXXXX
 RESULT 3
 ID US-07-938-990A-46 STANDARD PRI: 20 AA:
 XX XXXXX
 AC XXXXX
 DE Sequence 36, Application US/07938990A
 Sequence 36, Application US/07938990A
 GENERAL INFORMATION:
 CC APPLICANT: Griffith, Irwin J.
 CC APPLICANT: Pollock, Joanne
 CC APPLICANT: Bond, Julian
 CC TITLE OF INVENTION: Allergic Proteins And Peptides From
 CC NUMBER OF SEQUENCES:
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Lahive & Cockfield
 CC STREET: Sixty State Street
 CC CITY: Boston
 CC STATE: MA
 CC COUNTRY: USA
 CC ZIP: 02109
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC OPERATING SYSTEM: IBM PC compatible
 CC SOFTWARE: Patent Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/07/938,990A
 CC FILING DATE: 19920901
 CC CLASSIFICATION: 435
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 07/730,452
 CC FILING DATE: July 15, 1991
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 07/729,134
 CC FILING DATE: July 10, 1991
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Amy E. Mandragoras
 CC REGISTRATION NUMBER: 36,207
 CC REFERENCE/DOCKET NUMBER: IPC-0250C (IM1-028)
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (617) 227-7400
 CC TELEFAX: (617) 227-5941
 CC INFORMATION FOR SEQ ID NO: 36:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 20 amino acids
 CC TYPE: AMINO ACID

```

CC      TOPOLOGY: linear
CC      MOLECULE TYPE: peptide
CC      FRAGMENT TYPE: internal
CC      SEQUENCE 20 AA: 2080 MW: 2304 CN:

Query Match      93.0%  Score 53:  DB 4:  Length 20:
Best Local Similarity 100.0%:  Pred.No. 4.33e-01:
Matches      8:  Conservative  0:  Mismatches  0:  Indels  0:  Gaps  0:

DE      10 IKRVSNVI 17
QY      2 IKRVSNVI 9

RESULT      5
ID      US-08-467-023-36      STANDARD:      PRI: 20 AA:
XX      XX
AC      xxxxxx
XX      XX
DT      DT
XX      XX
DE      DE
XX      XX
CC      Sequence 36, Application US/08457023
CC      GENERAL INFORMATION:
CC      APPLICANT: Griffith, Irwin J.:
CC      APPLICANT: Pollock, Joanne:
CC      APPLICANT: Bond, Julian F.:
CC      APPLICANT: Garman, Richard D.:
CC      APPLICANT: Kuo, Mei-Chang:
CC      APPLICANT: Yeung, Siu-mei H.:
CC      APPLICANT: Brauer, Andrew:
CC      APPLICANT: Exley, Mark A.:
CC      APPLICANT: Powers, Steven P.:
CC      TITLE OF INVENTION: Allergic Proteins And Peptides From
CC      TITLE OF INVENTION: Japanese Cedar Pollen
CC      NUMBER OF SEQUENCES: 261
CC      CORRESPONDENCE ADDRESS:
CC      ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
CC      STREET: 610 Lincoln St
CC      CITY: Waltham
CC      STATE: MA
CC      COUNTRY: USA
CC      ZIP: 02154
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: PatentIn Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: US/08/467-023
CC      FILING DATE: June 6, 1995
CC      CLASSIFICATION: 424
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: US/950-425
CC      FILING DATE: December 17, 1994
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: Jane E. Remillard
CC      REGISTRATION NUMBER: 38,872
CC      REFERENCE/DOCKET NUMBER: 025.6 USD2 (IM2-028CP52)
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: (617) 227-7400
CC      TELEFAX: (617) 227-5941
CC      INFORMATION FOR SEQ ID NO: 36:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 20 amino acids
CC      TYPE: amino acid
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: peptide
CC      FRAGMENT TYPE: internal
CC      SEQUENCE 20 AA: 2080 MW: 2304 CN:

Query Match      93.0%  Score 53:  DB 10:  Length 20:
Best Local Similarity 100.0%:  Pred.No. 4.33e-01:
Matches      8:  Conservative  0:  Mismatches  0:  Indels  0:  Gaps  0:

DB      10 IKRVSNVI 17

```

```

17 2 KRVSNI 9
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
10
```

```

RESULT      7
XX US-08-467-697-36 STANDARD: PRT: 20 AA.
XX xxxxxx
XX
XX AC
XX XX
XX DI
XX DE
XX SE
XX Sequence 36, Application US/8467697
XX
XX Sequence 36, Application US/8467697
XX GENERAL INFORMATION:
XX APPLICANT: Griffith, J.W., Jr.
XX APPLICANT: Pollock, Joan
XX APPLICANT: Bond, Julian
XX APPLICANT: Gartner, Richard
XX APPLICANT: Kuo, Mei-Chia
XX APPLICANT: Yeung, Siu-mei
XX APPLICANT: Brauer, Andrew
XX APPLICANT: Exley, Mark A.
XX APPLICANT: Powers, Steven V.
XX TITLE OF INVENTION: Allergic Proteins And Peptides From
XX Japanese Cedar Pollen
XX NUMBER OF SEQUENCES: 251
XX CORRESPONDENCE ADDRESS:
XX ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
XX STREET: 610 Lincoln St
XX CITY: Waltham
XX STATE: MA
XX COUNTRY: USA
XX ZIP: 02154
XX
XX COMPUTER READABLE FORM:
XX MEDIUM TYPE: Floppy disk
XX COMPUTER: IBM PC Compatible
XX OPERATING SYSTEM: PC-DOS/MS-DOS
XX SOFTWARE: Patent In Release #1.0, Version #1.25
XX CURRENT APPLICATION DATA:
XX APPLICATION NUMBER: US/08/467,697
XX FILING DATE: June 6, 1995
XX CLASSIFICATION: 424
XX PRIOR APPLICATION DATA:
XX APPLICATION NUMBER: 387,60,225
XX FILING DATE: December 6, 1994
XX ATTORNEY/AGENT INFORMATION:
XX NAME: Jane E. Remilla
XX REGISTRATION NUMBER: 872
XX REFERENCE/SOCKET NUMBER: 025.6 USM (JMI-0280-FIN)
XX TELECOMMUNICATION LINE RX/TX CN:
XX TELEPHONE: (617) 427-6400
XX TELEFAX: (617) 427-6541
XX INFORMATION FOR SEQ ID NO: 36:
XX SEQUENCE CHARACTERISTICS:
XX LENGTH: 20 amino acids
XX TYPE: amino acid
XX TOPOLOGY: linear
XX MOLECULE TYPE: peptide
XX FRAGMENT TYPE: internal
SQ SEQUENCE 20 AA: 2080 MW: 2304 CN:
Query Match 93.0% Score 53; DB 10; Length 20;
Best Local Similarity 100.0%; Pred.No. 4.31e+01;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 10 IKRVSNVI 17
QY IIIIII
QY 2 IKRVSNV; 9
RESULT 8
ID US-09-142-524-4 STANDARD: PRT: 31 AA.
XX
XX AC xxxxxx

```

XX.
DT
XX
DE
XX
XX
Sequence 4, Application US/09142524
GENERAL INFORMATION:
APPLICANT: Sore, Toshio
APPLICANT: Kume, Akio
APPLICANT: Kairiki, Kazuo
APPLICANT: Iwama, Akiko
APPLICANT: Kico, Kohsuke
TITLE OF INVENTION: Peptide-based Immunotherapeutic Agent For Treating
FILE OF INVENTION: Allergic Diseases
FILE REFERENCE: Docket No. SPO-103
CURRENT APPLICATION NUMBER: US/09/142,524
CURRENT FILING DATE: 1999-01-04
EARLIER APPLICATION NUMBER: 8/80/702
EARLIER FILING DATE: 1996-03-10
EARLIER APPLICATION NUMBER: PCT/JP97/00740
EARLIER FILING DATE: 1997-03-10
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 31
TYPE: PRI
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Description of Unknown Organism:peptide
SEQUENCE 3: AA: 3661 MW: 4960 CN:
Query Match 93.0% Score 53: DB 17: Length 31:
Best Local Similarity 100.0% Pred. No. 4.33e+01:
Matches 8: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
DQ 24 IKRVSNV: 31
QY 2 IKRVSNV: 9
RESULT 9
ID US-09-142-524-5 STANDARD: PRI: 3: AA:
XX
XX
XX
XX
XX
Sequence 5, Application US/09142524
GENERAL INFORMATION:
APPLICANT: Sore, Toshio
APPLICANT: Kume, Akio
APPLICANT: Kairiki, Kazuo
APPLICANT: Iwama, Akiko
APPLICANT: Kico, Kohsuke
TITLE OF INVENTION: Peptide-based Immunotherapeutic Agent For Treating
FILE OF INVENTION: Allergic Diseases
FILE REFERENCE: Docket No. SPO-103
CURRENT APPLICATION NUMBER: US/09/142,524
CURRENT FILING DATE: 1999-01-04
EARLIER APPLICATION NUMBER: 8/80/702
EARLIER FILING DATE: 1996-03-10
EARLIER APPLICATION NUMBER: PCT/JP97/00740
EARLIER FILING DATE: 1997-03-10
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 31
TYPE: PRI
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Description of Unknown Organism:peptide

SQ SEQUENCE 3: AA: 3585 MW: 4902 CN:
Query Match 93.0% Score 53: DB 17: Length 31:
Best Local Similarity 100.0% Pred. No. 4.33e+01:
Matches 8: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
DQ 24 IKRVSNV: 31
QY 2 IKRVSNV: 9
RESULT 10
ID US-07-938-990A-62 STANDARD: PRI: 60 AA:
XX
XX
XX
XX
XX
Sequence 62, Application US/07938990A
GENERAL INFORMATION:
APPLICANT: Griffith, Irwin J.
APPLICANT: Pollock, Joanne
APPLICANT: Bond Julian
TITLE OF INVENTION: Allergenic Proteins And Peptides From
TITLE OF INVENTION: Japanese Cedar Pollen
NUMBER OF SEQUENCES: 70
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lahive & Cockfield
STREET: Sixty State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/938,990A
FILING DATE: 19920901
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 7/938,490
FILING DATE: July 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 7/938,494
FILING DATE: July 1992
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragouras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IPC-025CC (IMI-028)
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 62:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
SEQUENCE 60 AA: 6644 MW: 19464 CN:
Query Match 93.0% Score 53: DB 4: Length 60:
Best Local Similarity 100.0% Pred. No. 4.33e+01:
Matches 8: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
DQ 50 IKRVSNV: 57
QY 2 IKRVSNV: 9

RESULT 11
ID US-09-458-940-62 STANDARD: PRT: 60 AA
XX
AC xxxxxx
XX
DT
XX
XX
DE Sequence 62, Application US/08466940
XX
CC Sequence 62, Application US/08466940
CC GENERAL INFORMATION:
CC APPLICANT: Griffith, Irwin J.
CC APPLICANT: Pollock, Joanne
CC APPLICANT: Bond, Julian F.
CC APPLICANT: Garman, Richard O.
CC APPLICANT: Kuo, Mei-Chang
CC APPLICANT: Yeung, Siu-mei H.
CC APPLICANT: Brauer, Andrew
CC APPLICANT: Exley, Mark A.
CC APPLICANT: Powers, Steven P.
CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
CC TITLE OF INVENTION: Japanese Cedar Pollen
CC NUMBER OF SEQUENCES: 26
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
CC STREET: 610 Lincoln St
CC CITY: Waltham
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02154
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08466940
CC FILING DATE:
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/350,225
CC FILING DATE: December 6, 1994
CC APPLICATION NUMBER: 08/226,248
CC FILING DATE: April 9, 1994
CC APPLICATION NUMBER: 07/638,940
CC FILING DATE: September 1, 1992
CC APPLICATION NUMBER: PCT/US93/00139
CC FILING DATE: January 15, 1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Darlene A. Vassilopoulos
CC REGISTRATION NUMBER: 35,759
CC REFERENCE/AGENT INFORMATION:
CC TELEPHONE: (617) 466-6000
CC TELEFAX: (617) 466-6000
CC INFORMATION FOR SEQ ID NO: 62:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 60 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC FRAGMENT TYPE: internal
CC SEQUENCE 60 AA: 6644 MW: 19454 CN:
Query Match 93.0% Score 53; DB 10; Length 60;
Best Local Similarity 100.0%; Pred. No. 4,338-01;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 50 IKRVSNI: 57
QY 2 IKRVSNI: 9

RESULT 12
ID US-08-226-248A-62 STANDARD: PRT: 60 AA
XX
AC xxxxxx
XX
DT
XX
XX
DE Sequence 62, Application US/08226248A
XX
CC Sequence 62, Application US/08226248A
CC GENERAL INFORMATION:
CC APPLICANT: Griffith, Irwin J.
CC APPLICANT: Pollock, Joanne
CC APPLICANT: Bond, Julian F.
CC APPLICANT: Garman, Richard O.
CC APPLICANT: Kuo, Mei-Chang
CC APPLICANT: Yeung, Siu-mei H.
CC APPLICANT: Brauer, Andrew
CC APPLICANT: Exley, Mark A.
CC APPLICANT: Powers, Steven P.
CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
CC TITLE OF INVENTION: Japanese Cedar Pollen
CC NUMBER OF SEQUENCES: 26
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
CC STREET: 610 Lincoln St
CC CITY: Waltham
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02154
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: 08/226,248A
CC FILING DATE: April 9, 1994
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/358,790
CC FILING DATE: September 1, 1992
CC APPLICATION NUMBER: PCT/US93/00139
CC FILING DATE: January 15, 1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Darlene A. Vassilopoulos
CC REGISTRATION NUMBER: 35,759
CC REFERENCE/AGENT INFORMATION:
CC TELEPHONE: (617) 466-6000
CC TELEFAX: (617) 466-6000
CC INFORMATION FOR SEQ ID NO: 62:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 60 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC FRAGMENT TYPE: internal
CC SEQUENCE 60 AA: 6644 MW: 19454 CN:
Query Match 93.0% Score 53; DB 7; Length 60;
Best Local Similarity 100.0%; Pred. No. 4,338-01;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 50 IKRVSNI: 57
QY 2 IKRVSNI: 9

RESULT 13
ID US-08-467-697-62 STANDARD: PRT: 60 AA

CC APPLICANT: Garman, Richard D:
 CC APPLICANT: Kuo, Mei-Chang:
 CC APPLICANT: Yeung, Siu-mei H.:
 CC APPLICANT: Brauer, Andrew:
 CC APPLICANT: Exley, Mark A.:
 CC APPLICANT: Powers, Steven P.
 CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
 CC TITLE OF INVENTION: Japanese Cedar Pollen
 CC NUMBER OF SEQUENCES: 261
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
 CC STREET: 610 Lincoln St
 CC CITY: Waltham
 CC STATE: MA
 CC COUNTRY: USA
 CC ZIP: 02154
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC Compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: PatentIn Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/09/350,225
 CC FILING DATE: December 8, 1994
 CC CLASSIFICATION: 424
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/226,248
 CC FILING DATE: April 8, 1994
 CC APPLICATION NUMBER: 07/978,990
 CC FILING DATE: September 1, 1992
 CC APPLICATION NUMBER: PCI/US93/00139
 CC FILING DATE: January 25, 1993
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Darlene A. VanStoore
 CC REGISTRATION NUMBER: 35,729
 CC REFERENCE/DOCKET NUMBER: 025.6 US (IMI-028CPZ)
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (617) 466-6000
 CC TELEFAX: (617) 466-6040
 CC INFORMATION FOR SEQ ID NO: 62:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 60 amino acids
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: peptide
 CC FRAGMENT TYPE: internal
 CC SEQUENCE: 60 AA: 6644 MW: 19464 DA:

Query Match 94.0% Score 531 DB 81 Length 600
 Best Local Similarity 100.0% Prod. No. 4-93a-012
 Matches 61 Conservations 0 Mismatches 0 Gaps 0

Db 50 IKRVSNI 57
 1111111
 QY 2 IKRVSNI 9

Search completed: Mon Jun 19 16:29:08 2000
 Job time: 16 secs.

Release 3.1A John F. Collins, BioComputing Research Unit,
Copyright (C) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd
Msearch_pf protein - protein database search, using Smith-Waterman algorithm
Run on: Mon Jun 19 16:18:47 2000: MasPar time 4.95 Seconds
Tabular output not generated.
85,798 Million cell updates/sec

Title: >US-09-142-524A-14
Description: (1-9) from USC9.42524A.pcp
Perfect Score: 57
Sequence: 1: SIKRVSNVI 9
Scoring table: PAM 150
Gap 15

Searched: 142080 seqs, 47172406 residues
Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: pir62
1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 22.352; Variance 25.242; scale 0.886

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | GB | ID | Description | Pred. No. |
|------------|-------|-------------|--------|----|--------|-----------------------|-----------|
| 1 | 53 | 91.1 | 374 | 2 | JC2124 | major allergen Cry j | 3,776-01 |
| 2 | 53 | 91.1 | 374 | 2 | JC2124 | major allergen Cry j | 3,776-01 |
| 3 | 50 | 87.7 | 328 | 2 | S51539 | pectate lyase - Asper | 1,850-00 |
| 4 | 47 | 82.5 | 358 | 2 | S73776 | MG269 homolog Fli-Orf | 8,656-00 |
| 5 | 45 | 78.9 | 192 | 2 | E70414 | hypothetical protein | 2,328-01 |
| 6 | 45 | 78.9 | 569 | 2 | D56374 | type II secretion sys | 2,328-01 |
| 7 | 44 | 77.2 | 142 | 1 | F70437 | nucleoside-diphosphat | 3,768-01 |
| 8 | 44 | 77.2 | 169 | 1 | R3K75 | ribosomal protein S5 | 3,768-01 |
| 9 | 44 | 77.2 | 173 | 2 | S77486 | ribosomal protein S5 | 3,768-01 |
| 10 | 44 | 77.2 | 175 | 2 | R01215 | cyclic GMP-dependent | 3,768-01 |
| 11 | 44 | 77.2 | 195 | 2 | A64377 | fumarate hydratase (E | 3,768-01 |
| 12 | 44 | 77.2 | 256 | 2 | E70430 | hypothetical protein | 3,768-01 |
| 13 | 44 | 77.2 | 447 | 2 | S56256 | alpha-1,6-mannosyl-gl | 3,768-01 |
| 14 | 44 | 77.2 | 457 | 2 | G70216 | histidine--tRNA ligas | 3,768-01 |
| 15 | 44 | 77.2 | 653 | 2 | T03686 | probable NAD-ADP-rib | 3,768-01 |
| 16 | 44 | 77.2 | 720 | 2 | S50920 | hypothetical protein | 3,768-01 |
| 17 | 44 | 77.2 | 968 | 2 | B64385 | malt-glpr intergenic | 3,768-01 |
| 18 | 43 | 75.4 | 103 | 2 | A27859 | histone H4.1 - slime | 6,028-01 |
| 19 | 43 | 75.4 | 174 | 2 | S73219 | ribosomal protein S5, | 6,028-01 |
| 20 | 43 | 75.4 | 219 | 2 | H75290 | hypothetical protein | 6,028-01 |
| 21 | 43 | 75.4 | 220 | 2 | S70644 | probable ribosomal pr | 6,028-01 |
| 22 | 43 | 75.4 | 240 | 2 | G71227 | hypothetical protein | 6,028-01 |
| 23 | 43 | 75.4 | 244 | 2 | C75195 | molybdopterin-guanine | 6,028-01 |

| | | | | | | | |
|----|----|------|------|---|--------|------------------------|----------|
| 24 | 43 | 75.4 | 318 | 2 | S37225 | ribose-phosphate pyro | 6,028-01 |
| 25 | 43 | 75.4 | 377 | 2 | G59937 | lipopolysaccharide bi | 6,028-01 |
| 26 | 43 | 75.4 | 390 | 2 | S 87 | hypothetical protein | 6,028-01 |
| 27 | 43 | 75.4 | 431 | 2 | G 24 | probable aspartokinas | 6,028-01 |
| 28 | 43 | 75.4 | 459 | 2 | I 06 | hypothetical protein | 6,028-01 |
| 29 | 43 | 75.4 | 583 | 2 | I 14 | hypothetical protein | 6,028-01 |
| 30 | 43 | 75.4 | 784 | 1 | A1704 | endopeptidase La (EC | 6,028-01 |
| 31 | 43 | 75.4 | 785 | 1 | QGVZ8 | D5 protein - vacuolar | 6,028-01 |
| 32 | 43 | 75.4 | 785 | 2 | B16347 | ATP/GTP-binding prote | 6,028-01 |
| 33 | 43 | 75.4 | 785 | 2 | D42615 | CSR protein - vaccini | 6,028-01 |
| 34 | 43 | 75.4 | 792 | 2 | F05487 | potassium transport p | 6,028-01 |
| 35 | 42 | 73.7 | 256 | 2 | A21127 | beta-lactamase (EC 3. | 9,576-01 |
| 36 | 42 | 73.7 | 269 | 2 | A11259 | hypothetical protein | 9,576-01 |
| 37 | 42 | 73.7 | 313 | 2 | A1184 | lysophospholipase L2 | 9,576-01 |
| 38 | 42 | 73.7 | 313 | 1 | A7137 | hypothetical protein | 9,576-01 |
| 39 | 42 | 73.7 | 366 | 2 | G 167 | hypothetical protein | 9,576-01 |
| 40 | 42 | 73.7 | 433 | 2 | S 41 | keratin, type I, epim | 9,576-01 |
| 41 | 42 | 73.7 | 503 | 2 | F 83 | L-xylose transp: A | 9,576-01 |
| 42 | 42 | 73.7 | 854 | 2 | I 07 | glycoprotein B - Alice | 9,576-01 |
| 43 | 42 | 73.7 | 1054 | 2 | E 29 | probable isoicucyl-tr | 9,576-01 |
| 44 | 42 | 73.7 | 1177 | 2 | F0190 | chromosome segregatio | 9,576-01 |
| 45 | 42 | 73.7 | 1179 | 2 | F0190 | probable chromosome a | 9,576-01 |

ALIGNMENTS

RESULT 1
ENTRY JC2124 #1:0 complete
TITLE major allergen Cry j I precursor (clone pCCI-15) - Japanese cedar
ORGANISM Cedar
DATE #formal_name Cryptomeria japonica #common_name Japanese cedar
14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 26-Aug-1999
ACCESSIONS JC2124
REFERENCE JC2123
#authors Sone, T.; Komiyama, N.; Shimizu, K.; Kusakabe, T.; Morikubo, K.; Kiro, K.
#journal Biochem. Biophys. Res. Commun. (1994) 199:619-625
#title Cloning and sequencing of cDNA coding for Cry j I, a major allergen of Japanese cedar pollen.
#cross-references MGI:94183234
#accession JC2124
#molecule_type mRNA
#residues 1-374 #label: SON
#cross-references GB:D26545; NID:9493633; PID:0100087; PID:9493634
#experimental_source pollen
#note the authors described carbohydrate binding site for residue 2.

CLASSIFICATION aspartic family proteinase (A159)
KEYWORDS glycoprotein; proteinase
FEATURE 1-21
22-374
158.131.293.354
#length 374 #molecular-weight 40702 #cys-ksum 3692
#domain stand sequence #status predicted #label SON
#product major allergen Cry j I (clone pCCI-15) #status predicted #label MA1
#binding site carbohydrate (Asn) (covalent) #status predicted
Query Match 93.0% Score 53; DB 2; Length 374;
Best Local Similarity 100.0% Pred. No. 3,776-01;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 131 IKRVSNVI 138
|||||
QY 2 IKRVSNVI 9
RESULT 2
ENTRY JC2123 #type complete
TITLE major allergen Cr j I precursor (clone pCCI-2-2) - Japanese cedar
ORGANISM Cedar
DATE #formal_name Cryptomeria japonica #common_name Japanese cedar
14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change

```

ACCESSIONS      25-Aug-1999
REFERENCE       J021231 PC2056
AUTHORS        Sumi, T.; Komiyama, N.; Shimizu, K.; Kasakabe, T.; Morikubo,
               K.; Kido, K.
JOURNAL        Biochem Biophys Res Commun. (1994) 199:619-625
TITLE          Cloning and sequencing of cDNA coding for Cry j 1, a major
               allergen of Japanese cedar pollen.
CROSS-REFERENCES MIM:9418234
ACCESSION      J021231
**molecule_type** RNA
**residues** 1-374 **label** SON
**cross-references** GB:025544; NID:9494531; PID:dl06086; PID:q493632
**experimental_source** pollen
ACCESSION      PC2765
**molecule_type** protein
**residues** 22-5358 **label** 299-327:345-372 **label** S02
**note** the authors described carbohydrate binding site for
               residue 279
CLASSIFICATION  as per family pectate lyase LA:59
KEYWORDS        glycoprotein; pollen
FEATURE
1-21           *domain signal sequence *status predicted *label: SIGN
22-374         *product major allergen cry j 1 (clone PC01-2-2) *status
               predicted *label: MAIN
158-191:293-354 *binding site carbohydrate (Asn) (covalent) *status
               predicted
SUMMARY         *length 374 *molecular_weight 40645 *checksum 2920
               93.0% Score 53; DB 2; Length 374;
               Query Match 93.0% Score 53; DB 2; Length 374;
               Best Local Similarity 100.0%; Pred. No. 3,776-00;
               Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 131 IKRVSNVI 148
Cy 2 IKRVSNVI 9
RESULT 7
ENTRY 7
TITLE 7
AUTHORS 7
ORGANISM 7
DATE 7
ACCESSIONS      S01509
REFERENCE       H01 M.C. Whitehead, M.P. Clevehead, T.E. Dean, M.A.
               Goff, Genet. (1995) 27:142-149
JOURNAL        Genet. (1995) 27:142-149
TITLE          Sequence analysis of the Aspergillus nidulans pectate lyase
               gene and cloning of a cDNA for binding of pectate lyase to
               pectate. A recombinant of carbon catabolite repression.
CROSS-REFERENCES MIM:9418234
ACCESSION      S01509
**status** preliminary
**molecule_type** DNA
**residues** 1-326 **label** HCM
SUMMARY         *length 326 *molecular_weight 34565 *checksum 3698
               87.7% Score 50; DB 2; Length 326;
               Query Match 87.7% Score 50; DB 2; Length 326;
               Best Local Similarity 87.5%; Pred. No. 1,856-00;
               Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Db 111 IKRVSNVI 118
Cy 2 IKRVSNVI 9
RESULT 4
ENTRY 4
TITLE 4
AUTHORS 4
ORGANISM 4
VARIETY 4

```

```

DATE          27-Feb-1997 *sequence_revision 25-Apr-1997 *text_change
               17-Jul-1998
ACCESSIONS      S73776
REFERENCE       H01 M.C. Whitehead, M.P. Clevehead, T.E. Dean, M.A.
               Goff, Genet. (1995) 27:142-149
AUTHORS        Himmelreich, R.; Hilbert, H.; Plagens, H.; Fritsch, E.;
               B.C. Herrmann, K.
JOURNAL        Nucleic Acids Res. (1996) 24:4420-4449
TITLE          Complete sequence analysis of the genome of the bacterium
               Mycoplasma prev. n. n.
CROSS-REFERENCES MIM:9418234
ACCESSION      S73776
**status** preliminary; nucleic acid sequence not shown;
               translation not shown
**molecule_type** DNA
**residues** 1-358 **label** HIM
**cross-references** EMBL:AE000441; GB:000089; NID:0167414; PID:0167414
**note** the nucleotide sequence was submitted to the EMBL data
               library, November 1996
GENETICS
**genetic_code** S03
SUMMARY         *length 358 *molecular_weight 42617 *checksum 8572
               Query Match 82.5% Score 47; DB 2; Length 358;
               Best Local Similarity 77.8%; Pred. No. 8,568-00;
               Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 349 SIKRVSKL 357
Cy 1 SIKRVSNVI 9
RESULT 5
ENTRY 5
TITLE 5
AUTHORS 5
ORGANISM 5
DATE 5
ACCESSIONS      E70414
REFERENCE       Deckert, G.; Watten, P.V.; Gaasterland, T.; Yohari, M.;
               Lenox, A.; Graham, D.E.; Overbeck, R.J.; Speed, M.A.;
               Keller, M.; Anagnostou, M.; Hult, R.; Feldman, R.A.; Smith,
               J.M.; Olsen, G.J.; Swanson, R.V.
JOURNAL        Nature (1998) 392:153-158
TITLE          The complete genome of the hyperthermophilic bacterium
               Aquifex aeolicus
CROSS-REFERENCES MIM:9418234
ACCESSION      E70414
**status** preliminary; nucleic acid sequence not shown;
               translation not shown
**molecule_type** DNA
**residues** 1-358 **label** A1
**cross-references** GB:AE000441; NID:0167414; PID:0167414; PID:0167414
**experimental_source** strain: VHS
GENETICS
**gene** aq_1324
SUMMARY         *length 192 *molecular_weight 22076 *checksum 754
               Query Match 78.9% Score 45; DB 2; Length 192;
               Best Local Similarity 66.7%; Pred. No. 2,326-01;
               Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
Db 111 AKRVKNNV 119
Cy 1 SIKRVSNVI 9
RESULT 6
ENTRY 6
TITLE 6
AUTHORS 6
ORGANISM 6
DATE 6

```



```

#journal: DNA Res. (1996) 3:109-136
#title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp. PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions.
#cross-references: MIM:97061201
#accession: S77485
#status: nucleic acid sequence not shown; translation not shown
#molecule_type: DNA
#residues: 1-173 #label: KAN
#cross-references: EMBL:D92905; GB:AB001339; NID:q1652360;
#accession: P10N:BA017331; P10:di018066; P10:q1652411
#note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

GENETICS
#gene: rps5
#classification: #superfamily Escherichia coli ribosomal protein S5
#keywords: Protein biosynthesis; ribosome
#summary: #length: 173 #molecular-weight: 16241 #checksum: 674

Query Match 77.2% Score 44; DB 2; Length 173;
Best Local Similarity 62.5%; Pred. No. 3.76e+01;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 24 LKRVSNV 41
QY 2 LKRVSNV 9

RESULT 10
ENTRY 101215 #type: complete
TITLE Cyclic GMP-dependent protein kinase homolog F6N23.2 - Arabidopsis thaliana
#organism: #formal_name Arabidopsis thaliana #common_name mouse-ear cress
#date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 20-Sep-1999
#accessions: 101215
#reference: 101215
#authors: Beisel, G.
#submission: Submitted to the EMBL Data Library, April 1998
#description: The sequence of A. thaliana F6N23.
#accession: 101215
#status: preliminary; translated from GB/ YRI/UDBJ
#molecule_type: DNA
#residues: 1-175 #label: G51
#cross-references: EMBL:AF036919; NID:q3047100; P10:q3047101;
#accession: GSPB:GN00043; AISB:F6N23.2

GENETICS
#gene: AISB:F6N23.2
#introns: 6/10 37/10 61/20 87/10 104/3
#summary: #length: 176 #molecular-weight: 16241 #checksum: 3611

Query Match 77.2% Score 44; DB 2; Length 175;
Best Local Similarity 55.6%; Pred. No. 3.76e+01;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Db 2 LKRVSNV 31
QY 1 LKRVSNV 9

RESULT 11
ENTRY A64377 #type: complete
TITLE fumarate hydratase (EC 4.2.1.2) - Methanococcus jannaschii
#organism: #formal_name Methanococcus jannaschii
#date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 16-Jul-1999
#accessions: A64377
#reference: A64300
#authors: Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, J.A.; FitzGerald, L.M.; Clayton,

```

```

R.A.; Gocayne, J.D.; Kerlavage, A.R.; Dougherty, P.A.; Tomb, J.F.; Adams, M.D.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weissbrock, K.G.; Merrick, J.M.; Binkley, A.; Scott, J.L.; Geohagen, N.S.M.; Weidman, J.F.; Paulsen, O.; Nguyen, D.; Utterback, J.R.; Kelley, J.M.; Petersen, J.B.; Sadow, P.W.; Hanna, M.G.; Cotton, M.G.; Klotzel, K.M.; Hurst, M.A.; Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, R.O.; Woese, C.R.; Venturi, J.C. Science (1996) 273:1038-1073
#journal: Complete genome sequence of the methanogenic archaeon Methanococcus jannaschii.
#accession: A64377
#status: preliminary; nucleic acid sequence not shown; translation not shown
#molecule_type: DNA
#residues: 1-195 #label: BUL
#cross-references: GB:J67520; CR:137117; NID:q1591325; P10N:AA006121;
#accession: P10:q1591328; TIGR:MJ0617; P10:q151699

GENETICS
#map_position: F0R547001-547388
#start_codon: TIG
#classification: #superfamily iron-dependent tartrate dehydratase beta chain;
#keywords: iron-dependent tartrate dehydratase beta chain; hydrolase; carbon-oxygen lyase
#feature: 12-180
#summary: #domain iron-dependent tartrate dehydratase beta chain homology; #label: TIGP
#length: 195 #molecular-weight: 21846 #checksum: 6962

Query Match 77.2% Score 44; DB 2; Length 195;
Best Local Similarity 75.0%; Pred. No. 3.76e+01;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Db 136 SVKRYDNV 143
QY 1 SVKRYSNV 8

RESULT 12
ENTRY E70430 #type: complete
TITLE Hypothetical protein Aq1501 Aquifex aeolicus
#organism: #formal_name Aquifex aeolicus
#date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 08-May-1998
#accessions: E70430
#reference: A70300
#authors: Deckert, G.; Wallner, P.; Gaasterland, T.; Young, W.G.; Jorgensen, A.; Andersen, J.B.; Overbeek, R.; Shedd, M.A.; Kallstrom, A.; Birk, R.; Rasmussen, L.V.
#journal: Nature (1998) 393:161-165
#title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus
#cross-references: MIM:2419654
#accession: E70430
#status: preliminary; nucleic acid sequence not shown; translation not shown
#molecule_type: DNA
#residues: 1-256 #label: AUF
#cross-references: GB:AE007417; NID:q2983675; P10:q2983684; GB:AF005657
#experimental_source: strain VFS

GENETICS
#gene: aq_1501
#summary: #length: 256 #molecular-weight: 29513 #checksum: 2167

Query Match 77.2% Score 44; DB 2; Length 256;
Best Local Similarity 44.4%; Pred. No. 3.76e+01;
Matches 4; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Db 28 SVKRYGEV 36
QY 1 SVKRYSNV 9

```

```

**residues      1-457 **label KLE
**cross-references GB:AE 125; GB:AE000783; NID:g2688021;
                  PIDN:AE.56531.1; PID:g2688030; TIGR:BBO135
**experimental_source strain B31
CLASSIFICATION  *superfamily histidine--trna ligase; amino acid--trna ligase
                  *repeat homology; histidine--trna ligase homology
KEYWORDS         aminoacyl--trna synthetase; ligase; protein biosynthesis
FEATURE          8-418
SUMMARY          *domain histidine--trna ligase homology *label H1L
                  *length 457 *molecular-weight 52848 *checksum 4898

Query Match      77.2%; Score 44; DB 1; Length 457;
Best Local Similarity 56.7%; Pred. No. 3,76e-01;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 246 SIKRVSDVF 254
QY 1 SIKRVSNVI 9

RESULT 15
ENTRY   *type complete
TITLE   Probable NAD- Adp-ribosyltransferase (EC 2.4.2.30) - maize
ORGANISM *normal_name 2c. 19ys *common_name maize
DATE    24-Mar-1999 *sequence_revision 24-Mar-1999 *text_change
24-Mar-1999

ACCESSIONS  T03656
REFERENCE    214991
AUTHORS      Babyichuk, E.; Cottrill, P.; Storozhenko, S.; Fungation, M.;
              O'Farrell, M.; Van Montagu, M.; Inze, D.; Kushnir, S.
SUBMISSION  Submitted to the FBL Data Library, November 1997
DESCRIPTION Higher plants possess two poly(ADP-ribose) polymerases.
ACCESSION   T03656
STATUS      preliminary; translated from GB/EMBL/CDAS

**residues      1-653 **label BAB
**cross-references EMBL:AC222588; NID:e1264090; PID:e1264091;
                  PARP
GENETICS
GENE        DNA binding; glycosyltransferase; NAD: pentosyltransferase
KEYWORDS    *length 653 *molecular-weight 72995 *checksum 7074
SUMMARY

Query Match      77.2%; Score 44; DB 2; Length 653;
Best Local Similarity 62.5%; Pred. No. 3,76e-01;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 352 LKRSDVI 359
QY 2 IKRVSNI 9

Search completed: Mon Jun 19 16:18:53 2000
Job time : 6 secs.

```

```

S66256      *type complete
alpha-1,6-mannosyl-glycoprotein beta-1,
2-N-acetylglucosaminyltransferase (EC 2.4.1.143) - human
*normal_name Homo sapiens *common_name man
DATE        28-Oct-1995 *sequence_revision 13-Mar-1997 *text_change
07-May-1999

ACCESSIONS  S66256
REFERENCE    S66256
AUTHORS      Tan, J.; d'Ayostaro, G.A.F.; Bendjak, B.; Reck, F.; Sarkar,
              M.; Squire, J.A.; Leong, P.; Schachter, H.;
              Eur, J. Biochem. (1995) 231:317-328
              The human UDP-N-acetylglucosamine-6-phosphate-6-mannoside-beta-1,
              2-N-acetylglucosaminyltransferase II gene (MGAT2). Cloning
              of genomic DNA, localization to chromosome 14q21,
              expression in insect cells and purification of the
              recombinant protein.
**cross-references M01D:95351854
ACCESSION    S66256
**molecule_type DNA
**residues    1-447 **label TAN
**cross-references EMBL:U15128; NID:g302744; PID:g302745
10-28
59.86
FEATURE      *domain transmembrane *status predicted *label TMN
              *binding-site carbohydrate (Asn) (covalent) *status
              predicted
SUMMARY      *length 447 *molecular-weight 51550 *checksum 8407

Query Match      77.2%; Score 44; DB 2; Length 447;
Best Local Similarity 75.3%; Pred. No. 3,76e-01;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 53 GIKRVSNV 70
QY 1 SIKRVSNV 8

RESULT 14
ENTRY   *type complete
TITLE   Histidine--trna ligase (EC 6.1.1.21) his5 - Lyme disease
              Spirochaete
ALTERNATE_NAMES histidy--trna synthetase
ORGANISM *normal_name Borrelia burgdorferi *common_name Lyme disease
              Spirochaete
DATE      10-Sep-1999 *sequence_revision 10-Sep-1999 *text_change
10-Sep-1999

ACCESSIONS  G70116
REFERENCE    G70116
AUTHORS      Fraser, C.M.; Castens, S.; Huang, W.M.; Sutton, G.G.;
              Clayton, R.; Lathigra, R.; White, O.; Ketchum, K.A.;
              Dodson, R.; Hickey, E.K.; Gwinn, M.; Dougherty, B.; Tomb,
              J.F.; Fleischmann, R.D.; Richardson, D.; Peterson, J.;
              Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.;
              Vugt, R.V.; Palmer, N.; Adams, M.D.; Gocayne, J.; Weidman,
              J.; Uterback, L.; Watthey, L.; McDonald, L.; Artiach, P.;
              Bowman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst,
              K.; Roberts, K.; Hatch, B.; Smith, H.O.; Venter, J.C.
              Nature (1997) 390:580-586
              Genomic sequence of a Lyme disease spirochaete, Borrelia
              burgdorferi.
**cross-references M01D:98065943
ACCESSION    G70116
**status      preliminary; nucleic acid sequence not shown;
              translation not shown
**molecule_type DNA

```



Release 3.1A John F. Collins, BioComputing Research Unit,
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MFSrch_LPP protein - protein database search using Smith-Waterman algorithm

Run on: Mon Jun 19 16:17:57 2000: Maspar time 3.44 Seconds
Tabular output not generated. 79,741 Million cell updates/sec

Title: >US-09-142-524A-14
Description: (1-9) from US09142524A.pep
Perfect Score: 57
Sequence: 1 S:KRVSNV: 9

Scoring table:
PAM 150
Gap 15

Searched: 83857 seqs. 30454973 residues

Post-processing: Minimum Match 0%
Missing first 45 summaries

Database: swissprot38
Tr:swissprot

Statistics: Mean 22.859; Variance 21.707; scale 1.053

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARY:ES

| Result No. | Score | Query Match | Length | ID | Description | Pred. No. |
|------------|-------|-------------|--------|------------|------------------------|-----------|
| 1 | 53 | 94.0 | 324 | SRP_CRYJA | SUGI BASIC PROTEIN PRE | 4.46e-02 |
| 2 | 50 | 87.7 | 326 | PELENT | PERIATEINASE PRECURSO | 1.17e-01 |
| 3 | 47 | 82.5 | 358 | Y259_MYCPN | HYPOTHETICAL PROTEIN M | 1.57e-00 |
| 4 | 44 | 77.2 | 142 | NK_AQUAE | NUCLEOSIDE DIPHOSPHATE | 1.02e-01 |
| 5 | 44 | 77.2 | 169 | RSS_CVAPA | CYANELLE 7CS RIBOSOMAL | 1.02e-01 |
| 6 | 44 | 77.2 | 173 | RSS_SYNY3 | 30S RIBOSOMAL PROTEIN | 1.02e-01 |
| 7 | 44 | 77.2 | 195 | FUMB_METCA | PUTATIVE FUMARATE HYDR | 1.02e-01 |
| 8 | 44 | 77.2 | 447 | GNT2_HUMAN | ALPHA-1,6-MANNOSYLGLY | 1.02e-01 |
| 9 | 44 | 77.2 | 457 | SHR_BOREC | HISTIDYL-TRNA SYNTHETA | 1.02e-01 |
| 10 | 44 | 77.2 | 611 | MEN1_MOUSE | MENIN. | 1.02e-01 |
| 11 | 44 | 77.2 | 615 | MEN1_HUMAN | MENIN. | 1.02e-01 |
| 12 | 44 | 77.2 | 720 | YMS2_YEAS | HYPOTHETICAL 82.1 KD P | 1.02e-01 |
| 13 | 44 | 77.2 | 968 | Y582_METCA | HYPOTHETICAL PROTEIN M | 1.02e-01 |
| 14 | 43 | 75.4 | 102 | H4_PHYPO | HISTONE H4. | 1.75e-01 |
| 15 | 43 | 75.4 | 174 | R55_FORPU | CHLOROPLAST 30S RIBOSO | 1.75e-01 |
| 16 | 43 | 75.4 | 217 | R55_MYCLE | 30S RIBOSOMAL PROTEIN | 1.75e-01 |
| 17 | 43 | 75.4 | 220 | R55_MYCTU | 30S RIBOSOMAL PROTEIN | 1.75e-01 |
| 18 | 43 | 75.4 | 318 | KP2_YEAS | RIBOSE-PHOSPHATE PYROP | 1.75e-01 |
| 19 | 43 | 75.4 | 350 | APBE_SALLY | THIAMINE BIOSYNTHESIS | 1.75e-01 |
| 20 | 43 | 75.4 | 377 | Y21H_BACSU | HYPOTHETICAL 42.0 KD P | 1.75e-01 |
| 21 | 43 | 75.4 | 431 | AK_CHLIR | ASPARTOKINASE (EC 2.7. | 1.75e-01 |
| 22 | 43 | 75.4 | 730 | KOGA_MOUSE | DIACYLGLYCEROL KINASE, | 1.75e-01 |
| 23 | 43 | 75.4 | 784 | LON_RICPR | ATP-DEPENDENT PROTEASE | 1.75e-01 |

| | | | | | | | |
|----|----|------|------|---|------------|------------------------|----------|
| 24 | 43 | 75.4 | 785 | 1 | VD05_VACCV | PROTEIN D5. | 1.75e-01 |
| 25 | 43 | 75.4 | 785 | 1 | VD05_VARV | PROTEIN D5. | 1.75e-01 |
| 26 | 43 | 75.4 | 785 | 1 | VD05_VACCC | PROTEIN D5. | 1.75e-01 |
| 27 | 42 | 73.7 | 180 | 1 | RS5_SYNP6 | 30S RIBOSOMAL PROTEIN | 2.94e-01 |
| 28 | 42 | 73.7 | 256 | 1 | BLAB_BACCE | BETA-LACTAMASE, TYPE I | 2.94e-01 |
| 29 | 42 | 73.7 | 257 | 1 | BLAZ_BACSP | BETA-LACTAMASE, TYPE I | 2.94e-01 |
| 30 | 42 | 73.7 | 257 | 1 | BLAZ_BACCE | BETA-LACTAMASE, TYPE I | 2.94e-01 |
| 31 | 42 | 73.7 | 313 | 1 | BLAZ_BACCE | BETA-LACTAMASE, TYPE I | 2.94e-01 |
| 32 | 42 | 73.7 | 376 | 1 | BLAZ_BACCE | BETA-LACTAMASE, TYPE I | 2.94e-01 |
| 33 | 42 | 73.7 | 377 | 1 | BLAZ_BACCE | BETA-LACTAMASE, TYPE I | 2.94e-01 |
| 34 | 42 | 73.7 | 433 | 1 | BLAZ_BACCE | BETA-LACTAMASE, TYPE I | 2.94e-01 |
| 35 | 42 | 73.7 | 437 | 1 | BLAZ_BACCE | BETA-LACTAMASE, TYPE I | 2.94e-01 |
| 36 | 42 | 73.7 | 439 | 1 | BLAZ_BACCE | BETA-LACTAMASE, TYPE I | 2.94e-01 |
| 37 | 42 | 73.7 | 484 | 1 | BLAZ_BACCE | BETA-LACTAMASE, TYPE I | 2.94e-01 |
| 38 | 42 | 73.7 | 503 | 1 | BLAZ_BACCE | BETA-LACTAMASE, TYPE I | 2.94e-01 |
| 39 | 42 | 73.7 | 537 | 1 | BLAZ_BACCE | BETA-LACTAMASE, TYPE I | 2.94e-01 |
| 40 | 42 | 73.7 | 1374 | 1 | YAH8_SCHPO | HYPOTHETICAL HELICASE | 2.94e-01 |
| 41 | 41 | 71.9 | 561 | 1 | YQ34_YEAS | PUTATIVE ATP-DEPENDEN | 5.01e-01 |
| 42 | 41 | 71.9 | 897 | 1 | AP31_YEAS | AUTOPHAGY SERING/THREO | 5.01e-01 |
| 43 | 41 | 71.9 | 969 | 1 | MM22_MYCTU | PUTATIVE MEMBRANE PROT | 5.01e-01 |
| 44 | 41 | 71.9 | 996 | 1 | YH33_YEAS | HYPOTHETICAL 113.8 KC | 5.01e-01 |
| 45 | 41 | 71.9 | 1056 | 1 | YNN2_YEAS | HYPOTHETICAL 119.3 KC | 5.01e-01 |

A:IGNMENTS

RESULT: 1
ID SBP_CRYJA STANDARD: PRI: 374 AA.
AC P18632;
DT 01-NOV-1990 (Rel. 16, Createc
DT 01-NOV-1995 (Rel. 32, Last se
DE 15-DEC-1998 (Rel. 37, Last ab
DE SUGI BASIC PROTEIN PRECURSOR
OS Cryptomeria japonica (Japanes
OC Eukaryota; Viridiplantae; Str
OC eukaryophytes; Spermatophyta
OC Taxodiaceae; Cryptomeria.
RN [1]
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE
RC TISSUE+POLLEN;
RX MEDLINE: 94183234.
RA Sone T., Komiya N., Shimizu K., Kusakabe T., Morikubo K.,
RA Kino K.
RT Cloning and sequencing of cDNA coding for Cry J, a major allergen
of Japanese cedar pollen.
RL Biochem. Biophys. Res. Commun. 199:619-625(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE+POLLEN;
RA Nishida M., Kishimoto M., Kikuchi K., Kikuchi S., Kikuchi M.,
RA Nishida M. (1994) in: The EMBL/GenBank/DBJNC databases
RN [3]
RP SEQUENCE OF 22-41.
RC TISSUE+POLLEN;
RX MEDLINE: 89031257.
RA Tanai M., Ando S., Usui M., Kiritomo M., Sakaguchi M., Inouye S.,
RA Matubasi T.,
RT N-terminal amino acid sequence of a major allergen of Japanese cedar
pollen (Cry J I).
RL FEBS Lett. 239:329-332(1988).
RN [4]
RP CARBOHYDRATES
RX MEDLINE: 95003748.
RA Hijikata A., Matsumoto I., Kojima K., Ogawa H.,
RT Antigenicity of the oligosaccharide moiety of the Japanese cedar
(Cryptomeria japonica) pollen allergen, Cry J I.
RL Int. Arch. Allergy Immunol. 105:198-202(1994).
CC -!- PTM: CONTAINS FUCOSE/XULOSE-CONTAINING N-LINKED OLIGOSACCHARIDES.
CC -!- DISEASE: THIS PROTEIN IS THE MAJOR ALLERGEN OF JAPANESE CEDAR
POLLEN. THE MOST COMMON POLLEN ALLERGEN IN JAPAN.
CC -!- MISCELLANEOUS: THE SEQUENCE OF CRY J I FORM A IS SHOWN HERE. FORM
B DIFFERS IN SIX POSITIONS.
CC -!- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.

AMB A.1/AMB A.11/CRY J.1 SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: D26544: AAA05542.1: 1
 D26545: AAA05543.1: 1
 EMBL: D34699: AAA07020.1: 1
 PIR: A44773: A44773
 PFAM: PF00544: pec_lyase: 1
 PRINTS: PR00807: AMBALERGEN
 KW Allergen: Glycoprotein; Multigene family; Signal;
 FT SIGNAL 1 21
 FT CHAIN 22 374 SUGI BASIC PROTEIN.
 FT VARIANTE 12 12 L -> F (IN CRY J.1-B).
 FT VARIANTE 143 143 H -> Y (IN CRY J.1-B).
 FT VARIANTE 202 202 S -> T (IN CRY J.1-B).
 FT VARIANTE 221 221 L -> S (IN CRY J.1-B).
 FT VARIANTE 338 338 Q -> H (IN CRY J.1-B).
 FT VARIANTE 361 361 K -> Q (IN CRY J.1-B).
 FT VARIANTE 158 158 POTENTIAL.
 FT CARBOHYD 161 193 POTENTIAL.
 FT CARBOHYD 243 293 POTENTIAL.
 SEQUENCE 374 AA: 40645 MW: 74402595024856P CRC64:
 Query Match 93.08: Score 53: DB 1: Length 374:
 Best Local Similarity 100.00: Pred. No. 4,96e-02:
 Matches 8: Conservative C: Mismatches 0: Indels 0: Gaps 0:

Db 131 IKRVSNVI 138
 11111111
 QY 2 IKRVSNVI 9

RESULT 2 STANDARD: PRT: 426 AA.

ID PELEMENT
 AC Q00645: 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE PECTATE LYASE PRECURSOR (EN 4.2.2.2).
 GN PELA
 OS Emericella nidulans (Aspergillus nidulans).
 OC Eukaryota: Fungi: Ascomycota: Pleiomycetes: Emericellales
 OC Trichocommatales: Emericellaceae
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN: FGSC 4:
 RX MEDLINE: 95308556.
 RA HO M.C., Whitehead M.P., Cleveland T.E., Dear R.A.:
 RT "Sequence analysis of the Aspergillus nidulans pectate lyase gene and evidence for binding of promoter regions to CREA, a regulator of carbon catabolite repression.";
 RL Curr. Genet. 27:142-149(1995).
 CC 1- CATALYTIC ACTIVITY: ELIMINATIVE CLEAVAGE OF PECTATE TO GIVE GLUCOSACCHARIDES WITH 4-DEOXY-ALPHA-D-MANN-4-ENURONOSYL GROUPS AT THEIR NON-REDUCING ENDS.
 CC 1- SUBCELLULAR LOCATION: SECRETED (POTENTIAL).
 CC 1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: U05592: AAA80568.1: 1
 PFAM: PF00544: pec_lyase: 1
 KW Lyase; Signal;
 FT SIGNAL 1 15 POTENTIAL
 FT CHAIN 16 326 PECTATE LYASE.
 SO SEQUENCE 326 AA: 34580 MW: 76445A4A5D055049 CRC64:
 Query Match 87.78: Score 53: DB 1: Length 326:
 Best Local Similarity 97.58: Pred. No. 3.17e-01:
 Matches 7: Conservative C: Mismatches 0: Indels 0:

Db 111 IKRVSNVI 113
 11111111
 QY 2 IKRVSNVI 9

RESULT 3 STANDARD: PRT: 458 AA.

ID Y269_MYCPN
 AC P75395:
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE HYPOTHETICAL PROTEIN W3269 HXW.359.
 OS Mycoplasma pneumoniae.
 OC Bacteria: Firmicutes: Bacillus/clostridium group: Mollicutes.
 OC Mycoplasmataceae: Mycoplasma.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN: ATCC 29342 / M129.
 RX MEDLINE: 97165885.
 RA Himmelreich R., Hilbert H., Flanagan H., Parkl E., Li E.-C.
 RA Herrmann R.:
 RT "Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae.";
 RL Nucleic Acids Res. 24:4420-4449(1996).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: AE00044: AAB96038.1: 1
 KW Hypothetical protein.
 SO SEQUENCE 348 AA: 42617 MW: A1B80FA249E15D03 CRC64:
 Query Match 82.88: Score 42: DB 1: Length 348:
 Best Local Similarity 100.00: Pred. No. 1.77e-02:
 Matches 7: Conservative C: Mismatches 0: Indels 0:

Db 349 SIKRVSKLI 357
 11111111
 QY 1 SIKRVSNVI 9

RESULT 4 STANDARD: PRT: 142 AA.

ID NDK_AQUAE
 AC G67528:
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-FEB-2000 (Rel. 39, Last annotation update)
 DE NUCLEOSIDE DIPHOSPHATE KINASE (EC 2.7.4.6) (NDK) (NDP KINASE).
 GN NDK OR AQ.1590.
 OS Aquifex aeolicus
 OC Bacteria: Aquificales: Aquificaceae: Aquifex.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN: VF5:
 RX MEDLINE: 98196566.
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A., Graham D.E., Overbeek R., Sneed M.A., Keller M., Anjey M., Ruter P.

RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.:
 RI "The complete genome of the hyperthermophilic bacterium Aquifex
 RL aeolicus".
 RL Nature 392:353-358(1998).
 CC -1- FUNCTION: MAJOR ROLE IN THE SYNTHESIS OF NUCLEOSIDE TRIPHOSPHATES
 CC OTHER THAN ATP.
 CC -1- CATALYTIC ACTIVITY: ATP - NUCLEOSIDE DIPHOSPHATE - ADP -
 CC NUCLEOSIDE TRIPHOSPHATE.
 CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- SIMILARITY: BELONGS TO THE NDK FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AEO0746; AAC07481.1; .
 CC PROSITE: PS00463; NDP_KINASES; 1.
 CC PRAM: PF06334; NDK; 1.
 CC Transferase: Kinase; ATP-binding.
 CC ACT_SITE 120 120 BY SIMILARITY.
 CC SEQUENCE 142 AA: 189B50FA0317E998 CRC64:
 CC
 CC Query Match 77.2% Score 44: DR 1: Length 142:
 CC Best Local Similarity 55.6% Pred. No. 1.02e-01:
 CC Matches 5: Conservative 3: Mismatches 1: Indels 0: Gaps 0:
 CC
 CC Db 22 AIKRVREI 90
 CC :|||:|
 CC 1 SIKRVSNVI 9
 CC
 CC RESULT 5
 CC ID RS5_CVAPA STANDARD: PRT: 169 AA.
 CC AC P23402:
 CC DT 01-NOV-1997 (Rel. 20, Created)
 CC DI 01-NOV-1991 (Rel. 23, Last sequence update)
 CC DE 01-FEB-1996 (Rel. 33, Last annotation update)
 CC EMBL: The Cyanelle S-10 spec ribosomal protein gene operon from Cyanophora
 CC paradoxa.
 CC Mol. Gen. Genet. 224:222-231(1990).
 CC [2]
 CC SEQUENCE FROM N.A.
 CC STRAIN:LB555 / PRINGSHEIM:
 CC Stirewalt V.L., Michalowski C.B., Luffelhardt W., Bohnert H.J.,
 CC Bryant D.A.:
 CC Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases
 CC -1- SIMILARITY: BELONGS TO THE SSP FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M30487; AAA63629.1; .
 CC EMBL: J30921; AAA81219.1; .

DR PIR: S12220; R3KTS.
 DR HSSP: P02357; LPKP.
 DR MENDEL: 7927; CYAPA:tps5.1.
 DR PROSITE: PS00585; RIBOSOMAL_S5; 1.
 DR PRAM: PF00333; Ribosomal_S5; 1.
 KW Ribosomal protein; Cyanelle.
 SQ SEQUENCE 169 AA: 179B39D669812A9A CRC64:
 SQ
 CC Query Match 77.2% Score 44: DR 1: Length 169:
 CC Best Local Similarity 62.5% Pred. No. 1.02e-01:
 CC Matches 5: Conservative 3: Mismatches 0: Indels 0: Gaps 0:
 CC
 CC Db 24 IRRVSKV 31
 CC :|||:|
 CC 2 IKRVSNVI 9
 CC
 CC RESULT 6
 CC ID RS5_SYNV3 STANDARD: PRT: 173 AA.
 CC AC P71304:
 CC DT 01-NOV-1997 (Rel. 35, Created)
 CC DI 01-NOV-1997 (Rel. 35, Last sequence update)
 CC DE 01-NOV-1997 (Rel. 35, Last annotation update)
 CC EMBL: 30S RIBOSOMAL PROTEIN S5.
 CC GN RPSE OF RPS5 OR SLL1812.
 CC OS Synechocystis sp. (strain PC9803).
 CC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
 CC RN 11:
 CC SEQUENCE FROM N.A.
 CC MEDLINE: 97061201.
 CC RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 CC Miyajima T., Hiroseawa M., Sugitani M., Sasamoto S., Kimura T.,
 CC Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.,
 CC Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,
 CC Yamada M., Yasuda M., Tabata S.:
 CC "Sequence analysis of the genome of the unicellular cyanobacterium
 CC Synechocystis sp. strain PC9803. II. Sequence determination of the
 CC entire genome and assignment of potential protein-coding regions.".
 CC DNA Res. 3:109-136(1996).
 CC -1- FUNCTION: PROTEIN S5 IS IMPORTANT IN THE ASSEMBLY AND FUNCTION OF
 CC THE 30S RIBOSOMAL SUBUNIT (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SSP FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D90942; EAA17343.1; .
 CC HSSP: P02357; LPKP.
 CC PROSITE: PS00585; RIBOSOMAL_S5; 1.
 CC PRAM: PF00333; Ribosomal_S5; 1.
 KW Ribosomal protein.
 SQ SEQUENCE 173 AA: 16241 MW: A93C0F86E88475B1 CRC64:
 SQ
 CC Query Match 77.2% Score 44: DR 1: Length 173:
 CC Best Local Similarity 62.5% Pred. No. 1.02e-01:
 CC Matches 5: Conservative 3: Mismatches 0: Indels 0: Gaps 0:
 CC
 CC Db 24 IRRVSKV 31
 CC :|||:|
 CC 2 IKRVSNVI 9
 CC
 CC RESULT 7
 CC ID F0MB_METJA STANDARD: PRT: 195 AA.
 CC AC Q58034:
 CC DT 01-NOV-1997 (Rel. 35, Created)
 CC DI 01-NOV-1997 (Rel. 35, Last sequence update)
 CC DE 01-NOV-1997 (Rel. 35, Last annotation update)

RA Latnigra R., White C., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA Peterson J., Karlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA Van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA Utterback T., Watney L., McDonald L., Artach P., Bowman C.,
RA Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA Smith H.O., Venter J.C.,
RT "Genomic sequence of a Lyme disease spirochete, Borrelia burgdorferi,"
RL Nature 392:580-585(1992).
CC -:- CATALYTIC ACTIVITY: ATP -> L-HISTIDINE + TRNA(HIS) - ADP +
CC -:- PYROPHOSPHATE + L-HISTIDYL-TRNA(HIS).
CC -:- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -:- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -:- SIMILARITY: BELONGS TO CLASS-II AMINACYL-TRNA SYNTHETASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF001125; AAC6531.1;
DR EMBL: AF001125; AAC6531.1;
DR EMBL: AF016398; AAC79938.1;
KW Nuclear protein.
FT CONFLICT 457 457 M -> I (IN REF. 2).
FT CONFLICT 466 466 G -> E (IN REF. 2).
FT CONFLICT 512 512 I -> S (IN REF. 2).
SQ SEQUENCE 611 AA; 52849 MW; 1C6E9D5F8EF04988 CRC64;

Query Match 77.2% Score 44; DB 1; Length 457;
Best Local Similarity 65.7% Pred. No. 1.02e+01;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

DR 246 SIKRVSDVF 254
QY I I I I I I
QY I SIKRVSNVI 9

RESULT 10
ID MEN1_MOUSE STANDARD: PRI: 611 AA.
AC O88559;
DI 15-DEC-1998 (Rel. 37, Created);
DI 15-DEC-1998 (Rel. 37, Last sequence update);
DI 15-FEB-2000 (Rel. 39, Last annotation update);
DE MEN1.
GN Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Mus.
RN [1]
RC SEQUENCE FROM N.A.
RX STRAIN:129/GLA;
RA Bassett J.H.D., Thakker R.V.,
RT "Genomic sequence of mouse MEN1 gene,"
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE: 99039765.
RA Stewart C., Parente F., Piehl F., Farnebo F., Quincey D., Silins G.,
RA Bergman L., Carlo G.F., Lemmens I., Grimmond S., Khan C.Z., Khodel S.,
RA Teh B.T., Lagercrantz J., Sigters P., Calender A., van de Vem V.,
RA Kas K., Weber G., Hayward N., Gaudry P., Larsson C.,
RT "Characterization of the mouse Men1 gene and its expression during
RT development,"
RL Oncogene 17:2485-2493(1998).
CC -:- FUNCTION: NOT KNOWN.
CC -:- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -:- TISSUE SPECIFICITY: UBIQUITOUS. EXPRESSED AT HIGH LEVEL IN TESTIS
CC AND CNS.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF072755; AAC26001.1;
DR EMBL: AF072755; AAC26001.1;
DR EMBL: AF016398; AAC79938.1;
KW Nuclear protein.
FT CONFLICT 457 457 M -> I (IN REF. 2).
FT CONFLICT 466 466 G -> E (IN REF. 2).
FT CONFLICT 512 512 I -> S (IN REF. 2).
SQ SEQUENCE 611 AA; 67472 MW; C3FBB3A5655244E9 CRC64;

Query Match 77.2% Score 44; DB 1; Length 611;
Best Local Similarity 62.5% Pred. No. 1.02e+01;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

DB 118 VKKVSQVI 125
QY I I I I I I
QY 2 IRRVSNVI 9

RESULT 11
ID MEN1_HUMAN STANDARD: PRI: 615 AA.
AC O00255; O00632;
DI 15-DEC-1998 (Rel. 37, Created);
DI 15-DEC-1998 (Rel. 37, Last sequence update);
DI 15-FEB-2000 (Rel. 39, Last annotation update);
DE MEN1.
GN HOMO SAPIENS (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia;
OC Eutheria; Primates; Catarrhini; Hominoidea; Homo.
RN [1]
RC SEQUENCE FROM N.A. AND VARIANT S FMEN1.
RX TISSUE-LEUCOCYTE;
RX MEDLINE: 97258940.
RA Chandrasekharappa S.C., Guru D., Manickam P., Olufemi S.E.,
RA Collins F.S., Emmert-Buck M., Debelenko L.V., Zhuang Z.,
RA Lubensky I.A., Liotta L.A., Green J.S., Wang Y., Roe B.A.,
RA Weiseman J., Boquski M.S., Iwai S.K., Kester M.B., Kim Y.S.,
RA Heppner C., Dong Q., Spiegel A.M., Burns A.L., Marx S.J.,
RT "Positional cloning of the gene for multiple endocrine neoplasia-type
RT 1,"
RL Science 276:434-437(1997).
RN [2]
RN SEQUENCE FROM N.A. AND VARIANT S FMEN1.
RX MEDLINE: 97358593.
RA Agarwal S.K., Kester M.B., Debelenko L.V., Heppner C.,
RA Emmert-Buck M.R., Skarulis M.C., Doppman J.L., Kim Y.S.,
RA Lubensky I.A., Zhuang Z., Green J.S., Guru S.C., Manickam P.,
RA Olufemi S.E., Liotta L.A., Chandrasekharappa S.C., Collins F.S.,
RA Spiegel A.M., Burns A.L., Marx S.J.,
RT "Germline mutations of the MEN1 gene in familial multiple endocrine
RT neoplasia type 1 and related states,"
RL Hum. Mol. Genet. 6:1169-1175(1997).
RN [3]
RN VARIANT PARATHYROID ADENOMA LYS-26.
RX MEDLINE: 97385243.
RA Heppner C., Kester M.B., Agarwal S.K., Debelenko L.V.,
RA Emmert-Buck M.R., Guru S.C., Manickam P., Olufemi S.E.,
RA Skarulis M.C., Doppman J.L., Alexander R.H., Kim Y.S., Sagar S.K.,
RA Lubensky I.A., Zhuang Z., Liotta L.A., Chandrasekharappa S.C.,
RA Collins F.S., Spiegel A.M., Burns A.L., Marx S.J.,
RT "Somatic mutation of the MEN1 gene in parathyroid tumours,"
RL Nat. Genet. 16:375-378(1997).
RN [4]
RN VARIANT S FMEN1 D-42; P-165; D-169; S-188 AND E-289.
RX MEDLINE: 98130524.

RX MEDLINE: 84208772.
RA Wilhelm M., Wilhelm F., X.
RT "A transposon-like DNA fragment interrupts a Physarum polycephalum
RI histone H4 gene".
RL FEBS Lett. 155:249-254(1984).
CC -!- FUNCTION: HISTONE H4 ALONG WITH HISTONE H3, PLAYS A CENTRAL ROLE
CC IN NUCLEOSOME FORMATION.
CC -!- SUBUNIT: THE NUCLEOSOME IS AN OCTAMER CONTAINING TWO MOLECULES OF
CC H2A, H2B, H3, AND H4, WHICH WRAP APPROXIMATIVELY 146 BP OF DNA.
CC -!- MISCELLANEOUS: THE SEQUENCE OF H41 IS SHOWN.
CC -!- SIMILARITY: BELONGS TO THE HISTONE H4 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X00449; CAA25140.1;
CC EMBL: X15141; CAA33239.1;
CC EMBL: X15142; CAA33240.1;
CC EMBL: X0036; CAA68442.1;
CC PIR: A27859; A27859.
CC PIR: S10075; S10075.
CC PIR: S10076; S10076.
CC PRINTS: PR00623; HISTONEH4.
CC PROSITE: PS00047; HISTONE_H4; 1.
CC PFAM: PF00225; histone.
CC Chromosomal protein; Nucleosome core; Nuclear protein; DNA-binding.
FT INIT_MET 0
FT DNA_BIND 15 20
FT VARIANT 48 48 N -> K (IN H42).
FT SEQUENCE 102 AA; 1131; MW: 822379.5E72841FC CAG64;
CC
CC Query Match 75.4% Score 43; DB 1; Length 102;
CC Best Local Similarity 55.6% Pred. NO: 1.75e-01;
CC Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
CC
CC Db 42 GVRKISNT 50
CC 1 IKRVSNVI 9
CC
CC RESULT 15
CC ID RH5-PURP3 STANDARD: PRI: 174 AA.
CC AC P51238;
CC 01-06-1994 (Ref. 34, Unpubl.)
CC 01-06-1994 (Ref. 34, Last sequence update)
CC 01-06-1994 (Ref. 34, Last sequence update)
CC CHROMOFAST 305 HISTONE H4 PROTEIN SE
CC GN RP55.
CC OS Porphyra purpurea.
CC OC Chlorophyta.
CC OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.
CC RN 1;
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=AVONPORT;
CC RA Reith M.E., Munnholland J.;
CC RI "Complete nucleotide sequence of the Porphyra purpurea chloroplast
CC genome".
CC RL Plant Mol. Biol. Rep. 13:333-335(1995)
CC -!- SIMILARITY: BELONGS TO THE SSP FAMILY OF RIB-SOMAL PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

DR EMBL: U38804; AAC00184.1;
DR HSSP: P02357; IPKP.
DR MENDEL: 10355; PORPU:RPS5;1.
DR PROSITE: PS00585; RIBOSOMAL_S5; 1.
DR PFAM: PF00333; Ribosomal_S5; 1
KW Ribosomal protein; Chloroplast
SC SEQUENCE 174 AA; 18294 MW; 90528DA33252460C CAG64;
Query Match 75.4% Score 43; DB 1; Length 174;
Best Local Similarity 50.0% Pred. NO: 1.75e-01;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
Db 24 VKRVTKV 31
QY 2 IKRVSNVI 9
Search completed: Mon Jun 19 16:23:37 2000
Job time : 6 secs.

 W P E R E F

 (2)

Release 3.1A John F. Collins, Biocomputing Research Unit,
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution Rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Jun 19 16:18:20 2000: MasPar time 8.41 Seconds
 74.216 Million cell updates/sec

Tabular output not generated.

Title: >US-09-142-524A-14
 Description: (1-9) from US09.142524A.pep
 Perfect Score: 57
 Sequence: 1 SIKRVSNVI 9

Scoring table: PAM 150
 Gap 15

Searched: 225878 seqs, 69334:22 residues

Post-processing: Minimum Match 08
 Listing first 45 summaries

Database: seqrembl2
 1:sp_archea 2:sp_bacteria 3:sp_fungi 4:sp_human
 5:sp_invertebrate 6:sp_mammal 7:sp_mmc 8:sp_ordeaele
 9:sp-phage 10:sp_plant 11:sp_rodent 12:sp_unclassified
 13:sp-vertebrate 14:sp_virus

Statistics: Mean 21.880, Variance 21.726, scale 1.007

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description | Pred. No. |
|------------|-------|-------------|--------|----|--------------------------------|-----------|
| 1 | 49 | 86.0 | 868 | 3 | 073636 PHEROMONE RECEPTOR. | 1.04e+00 |
| 2 | 48 | 84.2 | 326 | 2 | 0924E5 PARA PROTEIN. | 1.87e+00 |
| 3 | 46 | 80.7 | 411 | 10 | 046322 PLASMIN GROWTH. COMPL | 5.90e+00 |
| 4 | 45 | 80.7 | 465 | 5 | 023157 W04D2.3 PROTEIN. | 5.90e+00 |
| 5 | 45 | 80.7 | 783 | 2 | 092439 BETA-N-ACETYLGLUCOSAMI | 5.90e+00 |
| 6 | 45 | 78.9 | 192 | 2 | 067344 HYPOTHETICAL 22.1 KD P | 1.03e+01 |
| 7 | 45 | 78.9 | 293 | 2 | 09X546 TETRACYCLINE REPRESSOR | 1.03e+01 |
| 8 | 45 | 78.9 | 360 | 2 | 092529 HYPOTHETICAL 41.5 KD P | 1.03e+01 |
| 9 | 45 | 78.9 | 569 | 1 | 029266 TYPE II SECRETION SYST. | 1.03e+01 |
| 10 | 45 | 78.9 | 682 | 3 | 013331 ACETOLACTATE SYNTHASE. | 1.03e+01 |
| 11 | 45 | 78.9 | 682 | 3 | 001154 ACETOLACTATE SYNTHASE. | 1.03e+01 |
| 12 | 44 | 77.2 | 175 | 10 | 065260 FEN2.2 PROTEIN. | 1.80e+01 |
| 13 | 44 | 77.2 | 256 | 2 | 067471 HYPOTHETICAL 29.9 KD P | 1.80e+01 |
| 14 | 44 | 77.2 | 433 | 10 | 092554 FISK9.17 PROTEIN. | 1.80e+01 |
| 15 | 44 | 77.2 | 434 | 10 | 092555 FISK9.16 PROTEIN. | 1.80e+01 |
| 16 | 44 | 77.2 | 512 | 2 | 044486 DINITROGENASE BETA SUB | 1.80e+01 |
| 17 | 44 | 77.2 | 610 | 11 | 09WVR8 MENIN. | 1.80e+01 |
| 18 | 44 | 77.2 | 653 | 10 | 050017 POLY(ADP-RIBOSE) POLYM | 1.80e+01 |
| 19 | 44 | 77.2 | 1098 | 4 | 015050 KIAAC342. | 1.80e+01 |
| 20 | 43 | 75.4 | 179 | 5 | 093715 F4339.2 PROTEIN. | 3.09e+01 |

21 43 75.4 242 1 074080 240RA LONG HYPOTHETICAL 3.09e+01
 22 43 75.4 251 2 094484 GALACTOSE-1-PHOSPHATE 3.09e+01
 23 43 75.4 293 2 091556 ABC TRANSPORTER ATP-B 3.09e+01
 24 43 75.4 308 10 065881 POPULUS TRICHOCARPA IS 3.09e+01
 25 43 75.4 347 5 091568 COS46.2. 3.09e+01
 26 43 75.4 390 2 059316 HYPOTHETICAL 43.0 KD P 3.09e+01
 27 43 75.4 431 2 084367 ASPARTOKINASE III 3.09e+01
 28 43 75.4 583 5 021472 SIMILAR TO CYTOCHROME 3.09e+01
 29 43 75.4 601 2 0940E1 ATP-DEPENDENT ZINC MET 3.09e+01
 30 43 75.4 785 14 093119 PUTATIVE 90.4K PROTEIN 3.09e+01
 31 43 75.4 792 10 064769 PUTATIVE POTASSIUM TRA 3.09e+01
 32 42 73.7 167 2 044026 HOXW AND ORF2 GENES 5.25e+01
 33 42 73.7 435 3 013352 MAP KINASE MINK. 5.25e+01
 34 42 73.7 435 2 094315 PXO-45. 5.25e+01
 35 42 73.7 511 2 094DK3 INITIATIVE PHASEASE. 5.25e+01
 36 42 73.7 854 14 036362 GLYCOPROTEIN B. 5.25e+01
 37 42 73.7 931 14 087544 104K PROTEIN. 5.25e+01
 38 42 73.7 1054 1 09VF67 1064AA LONG HYPOTHETICAL 5.25e+01
 39 42 73.7 1279 1 054462 1179AA LONG HYPOTHETICAL 5.25e+01
 40 42 73.7 1359 14 05W450 GKE MSV043 PUTATIVE RN 5.25e+01
 41 42 73.7 1984 6 028644 SODIUM CHANNEL ALPHA-S 5.25e+01
 42 42 73.7 2858 14 09ACE9 POLYPROTEIN. 5.25e+01
 43 42 73.7 4116 5 096958 CALO PROTEIN (EPASMENT 5.25e+01
 44 42 73.7 4120 5 09AYD1 PUSHOVER. 5.25e+01
 45 42 73.7 5322 5 09AYD2 PUSHOVER. 5.25e+01

ALIGNMENTS

RESULT 1
 ID 073636 PRELIMINARY: PRT: 868 AA.
 AC 073636;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DI 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DE 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
 DE PHEROMONE RECEPTOR.
 GN CA02.1.
 OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
 OC Neopterygii; Teleostei; Euteleostei; Acanthopterygii; Percomorphi;
 OC Tetraodontiformes; Tetraodontidae; Fugu.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 98226788.
 RA NAITO T., SAITO Y., YAMAMOTO J., NOZAKI Y., TOKURA K., HAZAKA M.,
 RA NAKANISHI S., BRENNER S.;
 RT "Putative pheromone receptors related to the Ca2+-sensing receptor in
 Fugu."
 RC Proc. Natl. Acad. Sci. U.S.A. 95:5128-5131(1998)
 RD EMSL: AB008654; BAA26123.1.
 DR PFAM: PF00093; Tm3.1.
 DR PFAM: PF01094; ANU_receptor_1.
 DR PRINTS: PR00248; GPCRAGR.
 KW Pheromone.
 SQ SEQUENCE 868 AA; 95531 MW; DA037852 CRC32;
 Query Match 86.08; Score 49; DR 13; Length 868;
 Best Local Similarity 75.08; Pred. No. 1.04e+00;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Db 282 IKRVAQVI 289
 QY 2 IKRVSNVI 9
 |||||

RESULT 2
 ID 0924E5 PRELIMINARY: PRT: 326 AA.
 AC 0924E5;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DI 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DE 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
 DE PARA PROTEIN.
 GN PARA.

OS Escherichia coli.
 CG Plasmid Incil Colib-p9.
 CC Bacteria: Proteobacteria, gamma subdivision; Enterobacteriaceae.
 CC Escherichia.
 RN [1].
 RP SEQUENCE FROM N.A.
 RA SAMPEL G., MIZUBUCHI K.:
 RI "Organization and diversification of plasmid genomes: complete
 RI nucleotide sequence of the Colib-p9 genome."
 RJ Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB021078; BAA76111;
 KW Plasmid.
 SQ SEQUENCE 326 AA: 36227 MW: 92469414 CRC32:

Query Match 84.2% Score 48: DB 2: Length 326:
 Best Local Similarity 75.0% Pred. NO. 1.87e-00:
 Matches 6: Conservative 2: Mismatches 0: Indels 0: Gaps 0:

DB 26: IKRVNVI 268
 IIII:II
 QY 2 IKRVNVI 9

RESULT 3 PRELIMINARY: PRI: 411 AA.
 ID Q46322
 AC Q46322
 DT 01-JUN-1998 (TRENBLrel. 06, Created)
 DI 01-JUN-1998 (TRENBLrel. 06, Last sequence update)
 DT 01-AUG-1998 (TRENBLrel. 07, Last annotation update)
 DE PLASMID GR04970, COMPLETE SEQUENCE.
 GN Gracillaria robusta.
 OS Gracillaria robusta.
 CC Plasmid GR04970.
 CC Eukaryota: Rhodophyta: Florideophyceae: Gracillariaceae:
 CC Gracillaria.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA GOFF L.J., KERN D.A.:
 RI Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 RJ EMBL: AF034718; AAC04724.1;
 KW Plasmid.
 SQ SEQUENCE 411 AA: 48427 MW: 6386593 CRC32:

Query Match 80.7% Score 46: DB 10: Length 411:
 Best Local Similarity 62.5% Pred. NO. 5.90e-00:
 Matches 5: Conservative 3: Mismatches 0: Indels 0: Gaps 0:

DB 10: IKRVNVI 17
 IIII:II
 QY 2 IKRVNVI 9

RESULT 4 PRELIMINARY: PRI: 465 AA.
 ID Q23157
 AC Q23157
 DT 01-NOV-1996 (TRENBLrel. 01, Created)
 DI 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
 DT 01-MAY-1999 (TRENBLrel. 10, Last annotation update)
 DE W04D2.3 PROTEIN.
 GN W04D2.3.
 OS Caenorhabditis elegans.
 CC Eukaryota: Metazoa: Nematoda: Secernentea: Rhabditia: Rhabditida:
 CC Rhabditina: Rhabditidae: Rhabditidae: Peloderinae: Caenorhabditis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA LENNARD N.:
 RI Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 RJ [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE: 94150718.
 RA WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M.,
 RA WILSON J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
 RA CRAXTON M., DEAR S., DU Z., DURSIN R., FAYELLO A., FULTON L.,
 RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,

RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
 RA LIGHTING J., LLOYD C., MCILRAY A., MORTIMORE B., O'CALLAGHAN M.,
 RA PARSONS J., PERCY C., RIFKEL L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
 RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT C., WOULDMAN P.,
 RT *2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans.*
 RL Nature 368:32-38(1994).
 DR EMBL: Z75552; CAA99940.1;
 DR PEARL: PF00876; Icnexin; 1.
 SQ SEQUENCE 465 AA: 54576 MW: 2FA84759 CRC32:

Query Match 80.7% Score 46: DB 5: Length 465:
 Best Local Similarity 100.0% Pred. NO. 5.90e-00:
 Matches 7: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

DB 456 KRVNVI 462
 IIII:II
 QY 3 KRVNVI 9

RESULT 5 PRELIMINARY: PRI: 78: AA.
 ID Q92H39
 AC Q92H39
 DT 01-MAY-1999 (TRENBLrel. 10, Created)
 DI 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
 DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)
 DE BETA-N-ACETYLGLUCOSAMINIDASE.
 GN CHIO.
 OS Pseudomonas sp. S9.
 CC Bacteria: Proteobacteria: gam. subdivision: Alteromonadaceae.
 CC Pseudomonas.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S9.
 RA TEKHARJANARUK S., GOODMAN A.E.:
 RI Cloning, sequence analysis and characterization of genes involved in
 RI chitin degradation of a marine bacterium, Pseudomonas sp. strain
 RI S9.*
 RL Microbiology 0:0(1999).
 DR EMBL: AF072374; AAC83237.1;
 DR HSSP: P06865; IQBC.
 SQ SEQUENCE 78: AA: 88856 MW: F887EDF8 CRC32:

Query Match 80.7% Score 46: DB 2: Length 78:
 Best Local Similarity 75.0% Pred. NO. 5.90e-00:
 Matches 6: Conservative 1: Mismatches 1: Indels 0: Gaps 0:

DB 392 IKRVNVI 194
 IIII:II
 QY 2 IKRVNVI 9

RESULT 6 PRELIMINARY: PRI: 192 AA.
 ID O67344
 AC O67344
 DT 01-AUG-1998 (TRENBLrel. 07, Created)
 DI 01-AUG-1998 (TRENBLrel. 07, Last sequence update)
 DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)
 DE HYPOTHETICAL 22.1 KD PROTEIN.
 GN AQ1324.
 OS Aquifex aeolicus.
 CC Bacteria: Aquificales: Aquificaceae: Aquifex.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VF5.
 RX MEDLINE: 98196666.
 RA DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
 RA GRAHAM D.E., OVERBEK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,
 RA FELDMAN R.A., SHORI J.M., OLSON G.J., SWANSON R.V.:
 RI *The complete genome of the hyperthermophilic bacterium Aquifex
 RI aeolicus.*

RL Nature 392:353-358(1998).

RM [2]

RP SEQUENCE FROM N.A.

RC STRAIN-VF5:

RA GRAHAM D.E., WARREN P.V., GAASSTERLAND T., YOUNG W.G., LENOX A.L.,

RA DECKER G., OVERBEK R., SNEAD M.A., KELLER M., AJJAY M., HUBER R.,

RA FEEDMAN R.A., SHORI J.M., OLSON G.J., SWANSON R.V.:

RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL: AE007341; AAC07306.1; -

KW Hypothetical protein.

SC SEQUENCE 192 AA: 22076 MW: 7528ADDD CRC32:

Query Match 78.9% Score 45: DB 2: Length 192:

Best Local Similarity 66.7% Pred. No. 1.03e+01:

Matches 6: Conservative 2: Mismatches 1: Indels 0: Gaps 0:

Db 111 AIKRVKNNV 119

QY 1 SIKRVSNVI 9

RESULT 7

ID Q9X546 PRELIMINARY: PRT: 293 AA.

AC Q9X546:

DI 01-NOV-1998 (Tremblrel. 12, Created)

DT 01-NOV-1998 (Tremblrel. 12, Last sequence update)

DE 01-NOV-1998 (Tremblrel. 12, Last annotation update)

GN TETRACYCLINE REPRESSOR PROTEIN HOMOLOG TRH.

OS Corynebacterium glutamicum.

OC Plasmid R-plasmid PAB2.

CC Bacteria: Firmicutes: Actinobacteria: Actinobacteridae:

OC Actinomycetales: Corynebacterineae: Corynebacteriaceae:

CC Corynebacterium.

EN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-22243:

RA TAUCH A., PFEHLER A., KALINOWSKI J.:

RT "DNA sequence and genetic organization of PAB2, a 19.8-kb R-plasmid of

RI Corynebacterium glutamicum encoding a new class of tetracycline efflux

RR and repressor proteins."

RL Submitted (JAN-1999) to the EMBL/GenBank/ DBJ databases.

DR EMBL: AF121000; AA025055.1; -

KW Cyclicin: Plasmid.

SC SEQUENCE 293 AA: 32350 MW: 5FLURAB3 CRC32:

Query Match 78.9% Score 45: DB 2: Length 293:

Best Local Similarity 55.6% Pred. No. 1.03e+01:

Matches 5: Conservative 4: Mismatches 0: Indels 0: Gaps 0:

Db 100 SMRRVANVV 108

QY 1 SIKRVSNVI 9

RESULT 8

ID Q9Z3L9 PRELIMINARY: PRT: 360 AA.

AC Q9Z3L9:

DI 01-MAY-1999 (Tremblrel. 10, Created)

DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)

DE 01-MAY-1999 (Tremblrel. 10, Last annotation update)

GN HYPOTHETICAL 41.5 KD PROTEIN.

OS Leptospira borgpetersenii.

CC Bacteria: Spirochaetales: Leptospiraceae: Leptospira.

EN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-L177:

RA KALANBAHETI T., BULACH D.M., RAJAKUMAR K., ADLER B.:

RT "Genetic Organization of the Lipopolysaccharide O-antigen Biosynthetic

RI Locus of Leptospira borgpetersenii Serovar Hardjibovis."

RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF078135; AA012950.1; -

KW Hypothetical protein.

SC SEQUENCE 360 AA: 41546 MW: 822C1EE3 CRC32:

Query Match 78.9% Score 45: DB 2: Length 360:

Best Local Similarity 75.0% Pred. No. 1.03e+01:

Matches 6: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

Db 129 IKRVSGVI 136

QY 2 IKRVSNVI 9

RESULT 9

ID C29266 PRELIMINARY: PRT: 569 AA.

AC C29266:

DI 01-JAN-1998 (Tremblrel. 05, Created)

DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)

DE 01-NOV-1998 (Tremblrel. 08, Last annotation update)

GN TYPE 11 SECRETION SYSTEM PROTEIN (OSPE-3).

OS Archaeoglobus fulgidus.

CC Archaea: Euryarchaeota: Archaeoglobales: Archaeoglobaceae:

CC Archaeoglobus.

EN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-VC-15 / DSM 4304 / ATCC 49558:

RX MEDLINE: 98C49343.

RA KLENK H.-P., CLAYTON R.A., TOMB J.-F., WHITE G., NELSON K.E.,

RA KETCHUM K.A., DODSON R.J., GWINN M., HICKEY E.K., PETERSON J.D.,

RA RICHARDSON D.L., KEAVACKENBUSH J., LEE N.H., SUTTON G.G., GILL S.,

RA FLEISCHMANN R.D., DOUGHERTY B.A., MCKENNEY K., ADAMS M.D., LORTUS S.,

RA KIRKNESS E.F., MCNEIL L.K., BADGER J.H., GLODEK A., ZHANG L.,

RA PETERSON S., REICH C.L., WEIDMAN J.F., McDONALD L., UFFERACK T.,

RA OVERBEK R., GOCCAYNE J.D., ARIACH P., KATNE B.P., SYKES S.M.,

RA COTTON M.D., SPRIGGS T., ROWMAN C., FUJII C., GARLAND S.M.,

RA SADOW P.W., D'ANDREA K.P., FRASER C.M., SMITH H.G., WOESE C.R.,

RA VENTER J.C.:

RT "The complete genome sequence of the hyperthermophilic, sulphate-

RI reducing archaeon Archaeoglobus fulgidus."

RL Nature 390:364-370(1997).

DR EMBL: AF001035; AA090245.1; -

DR TIGR: AF0996:

DR PFAM: PF00437; GSPIL_E: 1.

KW Hypothetical protein.

SC SEQUENCE 569 AA: 56520 MW: 6BF577EB CRC32:

Query Match 78.9% Score 45: DB 2: Length 569:

Best Local Similarity 75.0% Pred. No. 1.03e+01:

Matches 6: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

Db 544 IKRVSNVI 551

QY 2 IKRVSNVI 9

RESULT 10

ID O13331 PRELIMINARY: PRT: 682 AA.

AC O13331:

DI 01-JAN-1998 (Tremblrel. 05, Created)

DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)

DE 01-NOV-1999 (Tremblrel. 12, Last annotation update)

GN ACETOLACTATE SYNTHASE.

OS Magnaporthe oryzae (Rice blast fungus).

CC Eukaryota: Fungi: Ascomycota: Euascomycetes: Pyrenomycetes:

CC Phylacthales: Magnaportheaceae: Magnaporthe.

EN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-4091-5-8:

RA SWEIGARD J.A., CHUMLEY F.C., BROLL A.M., FARRALL L., VALENT B.:

RL Fungal Genetics Newsletter 44: 2-53(1997).

CC "SIMILARITY: WITH OTHER ENZYMES WHICH REQUIRE TPP.

DR EMBL: AF013601; AA01248.1; -

DR PROSITE: P500187; TPP_ENZYMES: 1.

DR PFAM: PF00205: TPP_enzymes: 1;
 KW Flavoproteinase: Thiamine pyrophosphatase;
 SQ SEQUENCE 682 AA: 73949 MW: 82F2886: CRC32:

Query Match 78.9% Score 45: DB 3: Length 682;
 Best Local Similarity 55.6% Pred. No. 1:03e-01;
 Matches 5: Conservative 4: Mismatches 0: Indels 0: Gaps 0:

DB 279 SIKRVADLV 287

QY 1 SIKRVSNVI 9

RESULT 11
 ID Q01154 PRELIMINARY: PRT: 682 AA:
 AC Q01154:
 DT 01-AUG-1998 (TrEMBLrel: 01, Created)
 DT 01-NOV-1998 (TrEMBLrel: 01, Last sequence update)
 DT 01-NOV-1999 (TrEMBLrel: 12, Last annotation update)
 DE ACETYLACTATE SYNTHASE;
 OS Magnaporthe oryzae (Rice blast fungus).
 CC Eukaryota; Fungi; Ascomycota; Euscomycetes; Pyrenomycetes;
 CC Phyllostales; Magnaporthaceae; Magnaporthe;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-6043:
 RA SWEIGARD J.A., FARRALL L., VALENT B.S.;
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBSJ databases.
 CC 1: SIMILARITY WITH OTHER ENZYMES WHICH REQUIRE TPP.
 DR EMBL: U39261; AA661669.1;
 DR PROSITE: PS0187; TPP_ENZYMES: 1;
 DR PFAM: PF00205: TPP_ENZYMES: 1;
 KW Flavoproteinase: Thiamine pyrophosphatase;
 SQ SEQUENCE 682 AA: 73873 MW: B5A3D9 CRC32:

Query Match 78.9% Score 45: DB 3: Length 682;
 Best Local Similarity 55.6% Pred. No. 1:03e-01;
 Matches 5: Conservative 4: Mismatches 0: Indels 0: Gaps 0:

DB 279 SIKRVADLV 287

QY 1 SIKRVSNVI 9

RESULT 12
 ID Q65260 PRELIMINARY: PRT: 175 AA:
 AC Q65260:
 DT 01-AUG-1998 (TrEMBLrel: 07, Created)
 DT 01-AUG-1998 (TrEMBLrel: 07, Last sequence update)
 DT 01-NOV-1999 (TrEMBLrel: 12, Last annotation update)
 DE PN2272 PROTEIN
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC eumhyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 CC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
 CC Arabidopsis;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, COLUMBIA;
 RA WASHU;
 RL Submitted (APR-1998) to the EMBL/GenBank/DBSJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, COLUMBIA;
 RA WATERSON R.;
 RL Submitted (APR-1998) to the EMBL/GenBank/DBSJ databases.
 DR EMBL: AF058913; AAC13612.1; "

DR PFAM: PF00027: CNMP_binding: 1;
 SQ SEQUENCE 175 AA: 19898 MW: 478978F9 CRC32:

Query Match 77.2% Score 44: DB 10: Length 175;
 Best Local Similarity 55.6% Pred. No. 1:80e-01;
 Matches 5: Conservative 4: Mismatches 0: Indels 0: Gaps 0:

DB 23 SCKRISLVV 21

QY 1 SIKRVSNVI 9

RESULT 13
 ID Q67471 PRELIMINARY: PRT: 256 AA:
 AC Q67471:
 DT 01-AUG-1998 (TrEMBLrel: 07, Created)
 DT 01-AUG-1998 (TrEMBLrel: 07, Last sequence update)
 DT 01-NOV-1998 (TrEMBLrel: 08, Last annotation update)
 DE HYPOTHELICAL 29.9 KD PROTEIN
 GN AQ1501.
 OS Aquilex aequilex.
 CC Bacteria; Aquificales; Aquificales; Aquilex;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VF5;
 RA DECKERT G., WARREN P.V., GAASERKAMP T., YOUNG W.G., LENOX A.L.;
 RA GRAHAM D.E., OVERBECK R., SNEY M.A., KELLER M., ADAY M., HUBER P.;
 RA FEDDAN R.A., SHORT J.M., GILSON G.J., SWANSON R.V.;
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBSJ databases.
 DR EMBL: AE00743; AAC07436.1; "
 KW Hypothetical protein;
 SQ SEQUENCE 256 AA: 29913 MW: E0588B0A CRC32:

Query Match 77.2% Score 44: DB 2: Length 256;
 Best Local Similarity 44.4% Pred. No. 1:80e-01;
 Matches 4: Conservative 5: Mismatches 0: Indels 0: Gaps 0:

DB 28 SVKPTGEVW 36

QY 1 SIKRVSNVI 9

RESULT 14
 ID Q62V54 PRELIMINARY: PRT: 143 AA:
 AC Q62V54:
 DT 01-MAY-1999 (TrEMBLrel: 10, Created)
 DT 01-MAY-1999 (TrEMBLrel: 10, Last sequence update)
 DT 01-NOV-1999 (TrEMBLrel: 12, Last annotation update)
 DE F15K9.17 PROTEIN.
 GN F15K9.17.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC eumhyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 CC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
 CC Arabidopsis;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, COLUMBIA;
 RA VYSOTSKAIA V.S., SCHWARTZ J.R., TORUMI M., YU G., LI J., LIU S.;
 RA KREMENTSKAYA I., LUROS J., ARAUJO R., BUEHLER E., CONWAY A.B.;
 RA DEWER K., FENG J., KIM C., LI Y., SHINN P., SUN H., DAVIS R.W.;
 RA ECKER J.R., FEDERSPIEL N.A., THEOLOGIS A.;
 RL Submitted (APR-1998) to the EMBL/GenBank/DBSJ databases.
 DR EMBL: AF058913; AAC13612.1; "

RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 RN 2.
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, COLUMBIA;
 RA THEOLOGIS A.;
 PL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 RN 3.
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, COLUMBIA;
 RA THEOLOGIS A.;
 PL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 RN 4.
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, COLUMBIA;
 RA THEOLOGIS;
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 RN EMBL: AC005278; AAC72119.1; .
 DR MENDEL: 33845; Arath:3000:33845.
 SQ SEQUENCE 433 AA: 45717 MW: 9331822E CRC32:

Query Match 77.2% Score 44 DB 10: Length 434;
 Best Local Similarity 75.0% Pred. No. 1.80e+01;
 Matches 6: Conservative 2: Mismatches 0: Indels 0: Gaps 0:

DL 317 SIKRVASV 324
 I I I I I I
 QY : SIKRVSNV 8

RESULT 15
 TO QZVSS PRELIMINARY: PRT: 434 AA.
 AC QZVSS.
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
 DE F15K9.16 PROTEIN.
 UN F15K9.16.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 CC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
 CC Arabidopsi.s.
 RN 11.
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, COLUMBIA;
 RA VYSOTSKAYA V.S., SCHWARTZ J.R., TORIUMI M., YU G., LI J., LIU S.,
 RA KREMERITSKAYA I., LUDOS J., ARAUJO R., BUEHLER E., CONWAY A.B.,
 RA DEMER K., FENG J., KIM C., LI Y., SHINN P., SUN H., DAVIS R.W.,
 RA ECKER J.R., FEDERSPEL N.A., THEOLOGIS A.;
 Arabidopsis thaliana chromosome 1; BAC F15K9 sequence.
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 RN 21.
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, COLUMBIA;
 RA THEOLOGIS A.;
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 RN 31.
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, COLUMBIA;
 RA THEOLOGIS A.;
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 RN 41.
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, COLUMBIA;
 RA THEOLOGIS;
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 RN EMBL: AC005278; AAC72120.1; .
 DR MENDEL: 33844; Arath:3000:33844.
 SQ SEQUENCE 434 AA: 46148 MW: 5ED90262 CRC32:

Query Match 77.2% Score 44 DB 10: Length 434;
 Best Local Similarity 75.0% Pred. No. 1.80e+01;
 Matches 6: Conservative 2: Mismatches 0: Indels 0: Gaps 0:

Db 318 SIKRVASV 325
 I I I I I I
 QY 1 SIKRVSNV 8

Search completed: Mon Jun 19 16:18:10 2000
 Job time : 10 secs.



(TM)

Release 3.1A John F. Collins, BioComputing Research Unit,
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MParch_pp protein - protein database search, using Smith-Waterman algorithm
Run On: Mon Jun 19 16:11:57 2000 MasPar time 3.71 Seconds
57,500 Million cell updates/sec

Tabular output not generated.

Title: >US-09-142-524A-14
Description: (1-9) from US09142524A pep
Perfect Score: 57
Sequence: 1 SIKRVSNVI 9

Scoring table: PAM 150
Gap 15

Searched: 148963 seqs, 23686106 residues

Post-processing: Minimum Watch 0%
Listing first 45 summaries
Maximum DB seq length 9

Database: arzenoseq35
japaneseeq

Statistics: Mean 15.426; Variance 44.241; scale 0.349

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result | No. | Score | Match | Length | DB ID | Description | Pred. No. |
|--------|-----|-------|-------|--------|--------|---------------------------------|-----------|
| 1 | 53 | 41.0 | 9 | 1 | W57751 | Residues 109-117 of Cr 1.87e-01 | |

Note: Post-processor removed 44 summaries from list due to search parameters chosen.

ALIGNMENTS

RESULT 1
ID W57751 standard: peptide: 9 AA.
AC W57751
DT 17-SEP-1998 (first entry)
DE Residues 109-117 of Cr 1.1.
KW Cr 1.1 Japanese cedar pollen antigen; allergy: immunotherapy;
KW HLA class II molecule.
OS Cryptomeria japonica.
PN W09823902-A1
PC 22-MAY-1998.
FE 12-NOV-1997; J04129.
PR 13-NOV-1996; JF-302053.
PA (MEIJ) MEIJ MILK PROD CO LTD.
FI Daiichi K. Kuro K. Kure A. Sone I.
DP WPI: 98-297517/25.

PT Peptides derived from Japanese cedar pollen antigens are
PT immunotherapeutic agents - useful for allergy treatment and typing
PT HLA class II molecules in allergy sufferers
PS Claim 12: Page 26: 50pp: Japanese.
CC This sequence represents residues 109-117 of the Cr 1.1 protein, and
CC is a peptide of the invention. The peptides are derived from Japanese
CC cedar pollen antigens, and are used as immunotherapeutic agents in the
CC treatment of allergy. The peptides can be used for identification and
CC typing of the particular HLA class II molecules in an allergy sufferer,
CC and also for peptide immunotherapy of an allergy. Using these peptides
CC the immunotherapy can be targeted more specifically to the requirements
CC of the individual patient, allowing more effective treatment of an
CC allergy, including those patients for whom treatment with a conventional
CC immunotherapeutic agent is ineffective.
SQ Sequence 9 AA:

Query Match 93.0%; Score 53; DB ID Length 9;
Best Local Similarity 100.0%; Pred. No. 1.87e-01;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 2 SIKRVSNVI 9
QY 2 SIKRVSNVI 9

Search completed: Mon Jun 19 16:12:04 2000
Job time: 7 secs.

.....
K E F E L A (TM)
.....

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MFarch_PP protein - protein database search, using Smith-Waterman algorithm
Run on: Tue Jun 20 13:39:35 2000 MasPar time 2.16 Seconds
60.035 Million cell updates/sec

Tabular output not generated.

File: US-09-142-524A-14
Description: (149) from USC9142524A.fpp
Perfect Score: 87
Sequence: 1 SKRVSNI: 9

Scoring table: PAM 150
Gap 15

Searched: 145141 seqs, 14437480 residues

Post-processing: Minimum Match 04
Listing first 45 summaries
Maximum DB seq length 9

Database: a-issued
1:5A_COMB 2:5B_COMB 3:6_COM 4:PCT_COMB 5:backfiles1

Statistics: Mean 14.393, Variance 44.418, scale 0.324

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARY

| Result | Query | No. | Score | Match | Length | DB | ID | Description | Pred. No. |
|--------|-------|-----|-------|-------|--------|----|----|-------------|-----------|
|--------|-------|-----|-------|-------|--------|----|----|-------------|-----------|

No matches found

Search completed: Tue Jun 20 13:39:38 2000
Job time : 3 secs.

No matches found
Search completed: Mon Jun 19 16:12:44 2000
Job time : 23 secs.

:*****

WIRELESS (TM)

Release 1.1A John F. Collins, BioComputing Research Unit,
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_PP protein : protein database search, using Smith-Waterman algorithm

Run on: Mon Jun 19 16:11:24 2000 Maspar time 3.89 Seconds
109.175 Million cell updates/sec

Tabular output not generated.

Title: US-09-142-524A-14
Description: (1-9) from US09142524A.pep
Perfect Score: 57
Sequence: : SIKKVSNI: 9

Scoring table: FAW 150
BAP 15

Searched: 142080 seqs, 47172406 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries
Maximum DB seq length 9

Database: pir62
pir1 2 pir2 3 pir3 4 pir4

Statistics: Mean 22.352; Variance 25.242; scale 0.886

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result | Query | No. | Score | Match | Length | DB | ID | Description | Pred. No. |
|--------|-------|-----|-------|-------|--------|----|----|-------------|-----------|
|--------|-------|-----|-------|-------|--------|----|----|-------------|-----------|

No matches found.

Search completed: Mon Jun 19 16:11:40 2000
Job time : 16 secs.

WIRE (TM)

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MFsrch_Ep protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Jun 19 16:10:32 2000 MasPar time 2.50 seconds
109.525 Million cell updates/sec

Tabular output not generated.

Title: >US-09-142-524A-14
Description: (1-9); from US09142524A.pep
Perfect Score: 57
Sequence: 1 S:KRVSNV1 9

Scoring table: PAM 150
Gap 15

Searched: 43857 seqs, 30454973 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries
Maximum DB seq length 9

Database: swiss-prot38
1:swissprot

Statistics: Mean 22.858; Variance 21.701; scale 1.053

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result | Rank | No. | Score | Match length (H L) | ID | Description | Pred. No. |
|--------|------|-----|-------|--------------------|----|-------------|-----------|
| | | | | | | | |

No matches found.

Search completed: Mon Jun 19 16:10:32 2000
Job time : 12 secs.

M E S S A G E (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit,
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Mpsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Jun 19 16:10:49 2000; MasPar time 5.97 Seconds
104.457 Million cell updates/sec

Tabular output not generated.

Title: US-09-142-524A-14
Description: (1-9) from US09142524A.ppt
Perfect Score: 57
Sequence: 1 SKKRVSNV: 9

Scoring table: PAM 150
Gap 15

Searched: 225878 seqs, 69334122 residues

Post-processing: Minimum Match: 0%
Listing first 45 summaries
Maximum DB seq length 9

Database: spiremb12
1:sp_archaea 2:sp_bacteria 3:sp_fungi 4:sp_human
5:sp_invertebrate 6:sp_mammal 7:sp_rbc 8:sp_organelle
9:sp_phage 10:sp_plant 11:sp_rodent 12:sp_unclassified
13:sp_vertebrate 14:sp_virus

Statistics: Mean 21.880; Variance 21.726; scale 1.007

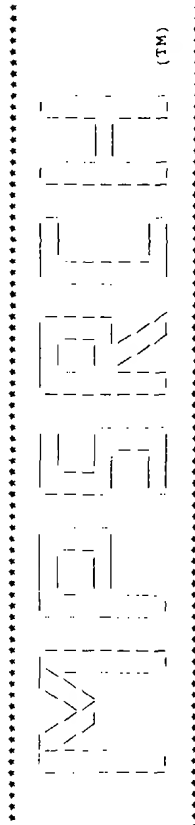
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result
No. | Query
Score | Match Length | DB ID | Description | Pred. No. |
|---------------|----------------|--------------|-------|-------------|-----------|
|---------------|----------------|--------------|-------|-------------|-----------|

NO matches found.

Search completed: Mon Jun 19 16:11:08 2000
Job time : 19 secs.



Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright: (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Msearch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Mon Jun 19 16:06:50 2000: Maspar time 3.66 Seconds
58,202 Million cell updates/sec
Tabular output not generated.

Title: US-09-142-524A-7
Description: (1-9) from US09142524A.rep
Perfect Score: 64
Sequence: 1 FIKRVSNVI 4

Scoring table: PAM 150
Gap 15

Searched: 188963 seqs, 23686106 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq35
l-geneseqp

Statistics: Mean 16.765, Variance 46.25, scale 0.362

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description | Pred. No. |
|------------|-------|-------------|--------|----------|--------------------------|-----------|
| 1 | 64 | 100.0 | 9 | 1 W57751 | Residues 109-117 of Cr | 1.25e+00 |
| 2 | 64 | 100.0 | 15 | 1 W57750 | Residues 106-120 of Cr | 1.25e+00 |
| 3 | 64 | 100.0 | 15 | 1 R6291 | Japanese cedar pollen | 1.25e+00 |
| 4 | 64 | 100.0 | 20 | 1 R45552 | Cry j 1: Pollen allergen | 1.25e+00 |
| 5 | 64 | 100.0 | 20 | 1 R82501 | Cry j 1: Japanese Cedar | 1.25e+00 |
| 6 | 64 | 100.0 | 30 | 1 W44683 | T-cell epitope peptide | 1.25e+00 |
| 7 | 64 | 100.0 | 31 | 1 W27372 | Multi-epitope peptide | 1.25e+00 |
| 8 | 64 | 100.0 | 31 | 1 W27373 | Multi-epitope peptide | 1.25e+00 |
| 9 | 64 | 100.0 | 80 | 1 W27369 | Multi-epitope peptide | 1.25e+00 |
| 10 | 64 | 100.0 | 105 | 1 W27370 | Multi-epitope peptide | 1.25e+00 |
| 11 | 64 | 100.0 | 134 | 1 W27371 | Multi-epitope peptide | 1.25e+00 |
| 12 | 64 | 100.0 | 353 | 1 R75388 | Japanese cedar pollen | 1.25e+00 |
| 13 | 64 | 100.0 | 353 | 1 R81587 | Cedar pollen allergen | 1.25e+00 |
| 14 | 64 | 100.0 | 374 | 1 R60166 | Japanese cedar pollen | 1.25e+00 |
| 15 | 64 | 100.0 | 374 | 1 R31937 | Cry j 1: | 1.25e+00 |
| 16 | 64 | 100.0 | 374 | 1 R82490 | Cry j 1: Japanese Cedar | 1.25e+00 |
| 17 | 64 | 100.0 | 374 | 1 R45541 | Cry j 1: Pollen allergen | 1.25e+00 |
| 18 | 49 | 76.6 | 292 | 1 R04895 | Penicillium-Insulin | 8.03e-01 |
| 19 | 48 | 75.0 | 367 | 1 R45577 | Jun s 1: | 1.05e-02 |
| 20 | 48 | 75.0 | 370 | 1 R45578 | Jun v 1: | 1.05e-02 |
| 21 | 47 | 73.4 | 424 | 1 W86150 | Human growth-related C | 1.36e-02 |
| 22 | 46 | 71.9 | 20 | 1 R45553 | Cry j 1: Pollen allergen | 1.77e-02 |
| 23 | 46 | 71.9 | 20 | 1 R82502 | Cry j 1: Japanese Cedar | 1.77e-02 |

ALIGNMENTS

RESULT 1
ID W57751 standard: peptide: 9 AA.
AC W57751:
DI 17-SEP-1998 (first entry)
DE Residues 109-117 of Cry j 1:
KW Cry j 1: Japanese cedar pollen antigen: allergy: immunotherapy:
KW HLA class II molecule:
OS Cryptomeria japonica:
PN W09820902-A1:
PD 22-MAY-1998:
PF 12-NOV-1997: J04129
PR 13-NOV-1996: JP-302053:
PA (MEIP) MEIJI MILK PROD CO LTD:
PI Daihiki K, Kino K, Kume A, Sone T:
DR WP: 98-297617/26:
PT Peptides derived from Japanese cedar pollen antigens are
PT immunotherapeutic agents - useful for allergy treatment and typing
PT HLA class II molecules in allergy sufferers
PS Claim 12: Page 24: 50pp: Japanese:
CC This sequence represents residues 109-117 of the Cry j 1 protein, and
CC is a peptide of the invention. The peptides are derived from Japanese
CC cedar pollen antigens, and are used as immunotherapeutic agents in the
CC treatment of allergy. The peptides can be used for identification and
CC typing of the particular HLA class II molecules in an allergy sufferer,
CC and also for peptide immunotherapy of an allergy. Using these peptides
CC the immunotherapy can be targeted more specifically to the requirements
CC of the individual patient, allowing more effective treatment of an
CC allergy, including those patients for whom treatment with a conventional
CC immunotherapeutic agent is ineffective.
SQ Sequence 9 AA:

Query Match 100.0%; Score 64; DB 1; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.25e+00;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Eb 1 FIKRVSNVI 9
Qy 1 FIKRVSNVI 9

RESULT 2
ID W57750 standard: peptide: 15 AA.
AC W57750:
DI 17-SEP-1998 (first entry)
DE Residues 106-120 of Cry j 1:
KW Cry j 1: Japanese cedar pollen antigen: allergy: immunotherapy:
KW HLA class II molecule:

Streptococcus pneumoniae 1.77e+02
Bankia goudii endogluc 1.77e+02
Bankia goudii endogluc 1.77e+02
Hexosaminidase enzyme 1.77e+02
Staphylococcus aureus 1.77e+02
Hepatitis C virus env 2.29e+02
Hepatitis C virus NS5 2.29e+02
Mouse synaptosomal class 2.29e+02
Mouse SNAP-25 polypept 2.29e+02
Synaptosomal associate 2.29e+02
Human thoracic aorta G 2.29e+02
Human thoracic aorta G 2.29e+02
PsbF (Wbpf) protein in 2.29e+02
Human multiple endocr 2.29e+02
Thermostable pyruvic a 2.29e+02
Tumor suppressor prote 2.29e+02
Staphylococcus aureus 2.29e+02
Staphylococcus aureus 2.29e+02
Adenylyate cyclase p101 2.29e+02
Synthetic anti-neoplas 2.29e+02
Peptide enhancer of 2.29e+02
Amphipathic peptide 2.29e+02

Cryptomeria japonica.
WO9820502-A1 (first entry)
12-MAY-1996: JP-302053.
13-NOV-1996: JP-302053.
(MEIP) MEIJU MILK PROD CO LTD.
Dairiki K, Kino K, Kume A, Sone T.
WPI: 98-23767/26.
Peptides derived from Japanese cedar pollen antigens are immunotherapeutic agents - useful for allergy treatment and typing HLA class II molecules in allergy sufferers
Claim 12: Page 26: Supp. Japanese.
This sequence represents residues 106-120 of the Cry j 1 protein, and is a peptide of the invention. The peptides are derived from Japanese cedar pollen antigens, and are used as immunotherapeutic agents in the treatment of allergy. The peptides can be used for identification and typing of the particular HLA class II molecules in an allergy sufferer, and also for peptide immunotherapy of an allergy. Using these peptides of the invention, a patient can be targeted more specifically to the requirements of the individual patient, allowing more effective treatment of an allergy, including those patients for whom treatment with a conventional immunotherapeutic agent is ineffective.
Sequence: 15 AA:

Query Match: 100.0%; Score 64; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.25e+00;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

4 FIKRVSNVI 12
|||||
1 FIKRVSNVI 9

R89291 standard; peptide: 15 AA.

12-MAR-1996 (first entry)
Japanese cedar pollen allergen Cryj 1 derived T-cell epitope peptide.
Japanese cedar pollen allergen; Cryj 1; T-cell epitope; peptides.
Prevention; treatment; cryptomeria pollenosis.
Cryptomeria japonica.
J07128295-A.
09-MAY-1995: 262626.
20-OCT-1993: JP-262626.
(MEIP) MEIJU MILK PROD CO LTD.
WPI: 95-20384/27.
New cryptomeria pollen allergen T cell epitope peptide - used for prevention, treatment and investigation of Japanese cedar pollenosis
Claim 5: Page 2: 8pp: Japanese.
R75388 is the Japanese cedar pollen allergen Cryj 1, from which the T-cell epitope peptides R89289-R89295 were derived. The peptides can be used for the prevention and treatment of cryptomeria pollenosis, and also for the investigation of pollenosis.
Sequence: 15 AA:

Query Match: 100.0%; Score 64; DB 1; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.25e+00;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

4 FIKRVSNVI 12
|||||
1 FIKRVSNVI 9

R45552 standard; Protein: 20 AA.

13-JUL-1994 (first entry)
Cry j 1 pollen allergen peptide CJI-11.
Japanese cedar; detection; allergy; treatment; diagnosis; T cell epitope; sensitivity.
Cryptomeria japonica.

WO9401560-A.
20-JAN-1994: U00139.
15-JAN-1993: WO-005561.
10-JUL-1992: US-538950.
01-SEP-1992: US-538950.
(IMMU-) IMMUNOLOGIC PHARM CORP.
Bond JF, Garman RD, Griffith IJ, Kuo M, Pollock Z.
WPI: 94-035055/04.
Antigens derived from Japanese cedar pollen allergen; Cry j 1; contain at least two T cell epitope(s); used to treat or diagnose allergy
Claim 1: Fig 13: 137pp: English.
The sequence is that of an isolated peptide of the Japanese cedar pollen allergen Cry j 1 (amino acids 101-120). The peptide CJI-11, can be used for the treatment and diagnosis of allergies associated with Japanese cedar pollen. It has enhanced therapeutic properties CC but reduced side effects compared to naturally occurring allergens SQ
Sequence: 20 AA:

Query Match: 100.0%; Score 64; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.25e+00;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

9 FIKRVSNVI 17
|||||
1 FIKRVSNVI 9

R82501 standard; Protein: 20 AA.
15-APR-1996 (first entry)
Cry j 1 Japanese Cedar pollen allergen; peptide fragment CJI-11.
Cry j 1; Japanese cedar pollen allergen; modified; drug production; allergy; Cryptomeria japonica.
Cryptomeria japonica.
WO9527786-A1.
19-OCT-1995:
06-APR-1995: U04245.
08-APR-1994: US-226248.
06-DEC-1994: US-350225.
(IMMU-) IMMUNOLOGIC PHARM CORP.
Chen X, Evans S, Franzen HM, Kuo M, Powers SP.
Shaked Z.
WPI: 95-36639/47.
Modified cryptomeria japonica (Cry j) 1 peptide(s) - useful for treating allergy to Japanese cedar pollen allergen or immunologically cross reactive allergens
Disclosure: Figure 2: 6pp: English.

Novel peptides of Cry j 1 have been modified as a part of a preformulation scheme to develop an optimised drug product for therapeutic treatment of humans suffering from allergy to Japanese cedar pollen allergen or an allergen which is immunologically cross reactive with Japanese cedar pollen allergen. Such modified peptides possess certain characteristics which render them particularly suitable for drug product formulation. Peptide fragments of Cry j 1 modified and unmodified, are given in R82491-R82525. This peptide fragment corresponds to amino acids 101-120 of the allergen mature protein.
Sequence: 20 AA:

Query Match: 100.0%; Score 64; DB 1; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.25e+00;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

9 FIKRVSNVI 17
|||||
1 FIKRVSNVI 9

R45552 standard; peptide: 30 AA.

W44683

DI 01-MAY-1998 (first entry)
DE Multi-epitope peptide used as immunotherapeutic agent #5.
KW T-cell epitope peptide; sugi pollen antigen; sugi pollinosis.
OS Synthetic.
PN Cryptomeria japonica.
PC J1002700-A.
PF 13-JAN-1998.
PR 24-JUN-1996: 163287.
PA (DAICEL) DAICEL CHEM IND LTD.
PA (WEIP) MEIJI SEIKA KAISHA LTD.
DR WPI: 98-13363/13.
PT T cell epitope peptide of sugi pollen antigen - used in the treatment of sugi pollinosis.
PS Claim 1: Page 4: 34pp: Japanese.
CC T-cell epitope peptides W4682-88 and their derivatives react with sugi pollinosis patient peripheral blood T lymphocytes. A composition prepared by combining at least 2 of the above peptides and/or their derivatives is used for the prevention and treatment of sugi pollinosis.
SQ Sequence 30 AA:
Query Match: 100.0%; Score 64; DB 1; Length 30;
Best Local Similarity 100.0%; Pred. No. 1.25e+00;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DI 9 FIKRVSNVI 17
:|||||
QY 1 FIKRVSNVI 9

RESULT 7
ID W27372 standard: peptide: 11 AA.
AC W27372
DI 24-MAR-1998 (first entry)
DE Multi-epitope peptide used as immunotherapeutic agent #4.
KW Multi-epitope peptide; immunotherapeutic agent; allergic disease.
KW T-cell epitope region; allergen; lymphocyte; immunoglobulin E.
OS Synthetic.
PN W09732600-A1.
PC 12-SEP-1997.
PR 10-MAR-1996: JP-080702.
PA (WEIP) MEIJI MILK PROD CO LTD.
PA Dairiki K, Iwama A, Kino K, Kume A, Some T.
DR WPI: 97-470495/43.
PT Peptide immunotherapeutic agent to treat allergic diseases - contains multi-epitope peptide containing T cell epitope regions from different allergens.
PS Claim 9: Page 32: 58pp: Japanese.
CC The present sequence represents a multi-epitope peptide which is used as a new immunotherapeutic agent. It comprises T cell epitope regions from 2 or more different allergens (preferably linked via arginine or lysine dimers), where the T cell epitope regions have a positivity index greater than 100 as measured in a patient group responding to the allergen; have at least 70% reactivity with lymphocytes from patients responding to the allergen; and are not reactive with immunoglobulin E (IgE) antibodies from patients responsive to the allergen. The agent can be used to prevent and treat a wide variety of allergic diseases, e.g. by desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
SQ Sequence 31 AA:
Query Match: 100.0%; Score 64; DB 1; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.25e+00;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DI 23 FIKRVSNVI 31
:|||||
QY 1 FIKRVSNVI 9

RESULT 8
ID W27373 standard: peptide: 31 AA.
AC W27373
DI 24-MAR-1998 (first entry)
DE Multi-epitope peptide used as immunotherapeutic agent #1.
KW Multi-epitope peptide; immunotherapeutic agent; allergic disease.
KW T-cell epitope region; allergen; lymphocyte; immunoglobulin E.
OS Synthetic.
PN W09732600-A1.
PC 12-SEP-1997.
PR 10-MAR-1996: JP-0807-2.
PA (WEIP) MEIJI MILK PROD CO LTD.
PA Dairiki K, Iwama A, Kino K, Kume A, Some T.
DR WPI: 97-470495/43.
PT Peptide immunotherapeutic agent to treat allergic diseases - contains multi-epitope peptide containing T cell epitope regions from different allergens.
PS Claim 6: Page 31: 58pp: Japanese.
CC The present sequence represents a multi-epitope peptide which is used as a new immunotherapeutic agent. It comprises T cell epitope regions from 2 or more different allergens (preferably linked via arginine or lysine dimers), where the T cell epitope regions have a positivity index greater than 100 as measured in a patient group responding to the allergen; have at least 70% reactivity with lymphocytes from patients responding to the allergen; and are not reactive with immunoglobulin E (IgE) antibodies from patients responsive to the allergen. The agent can be used to prevent and treat a wide variety of allergic diseases, e.g. by desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
SQ Sequence 31 AA:
Query Match: 100.0%; Score 64; DB 1; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.25e+00;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DI 23 FIKRVSNVI 31
:|||||
QY 1 FIKRVSNVI 9

RESULT 9
ID W27369 standard: peptide: AA.
AC W27369
DI 24-MAR-1998 (first entry)
DE Multi-epitope peptide used as immunotherapeutic agent #1.
KW Multi-epitope peptide; immunotherapeutic agent; allergic disease.
KW T-cell epitope region; allergen; lymphocyte; immunoglobulin E.
OS Synthetic.
PN W09732600-A1.
PC 12-SEP-1997.
PR 10-MAR-1996: JP-0807-2.
PA (WEIP) MEIJI MILK PROD CO LTD.
PA Dairiki K, Iwama A, Kino K, Kume A, Some T.
DR WPI: 97-470495/43.
PT Peptide immunotherapeutic agent to treat allergic diseases - contains multi-epitope peptide containing T cell epitope regions from different allergens.
PS Claim 6: Page 31: 58pp: Japanese.
CC The present sequence represents a multi-epitope peptide which is used as a new immunotherapeutic agent. It comprises T cell epitope regions from 2 or more different allergens (preferably linked via arginine or lysine dimers), where the T cell epitope regions have a positivity index greater than 100 as measured in a patient group responding to the allergen; have at least 70% reactivity with lymphocytes from patients responding to the allergen; and are not reactive with immunoglobulin E (IgE) antibodies from patients responsive to the allergen. The agent can be used to prevent and treat a wide variety of allergic diseases, e.g. by desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
SQ Sequence 80 AA:
Query Match: 100.0%; Score 64; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 1.25e+00;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DI 17 FIKRVSNVI 25

DI 24-MAR-1998 (first entry)
DE Multi-epitope peptide used as immunotherapeutic agent #5.
KW Multi-epitope peptide; immunotherapeutic agent; allergic disease.
KW T-cell epitope region; allergen; lymphocyte; immunoglobulin E.
OS Synthetic.
PN W09732600-A1.
PC 12-SEP-1997.
PR 10-MAR-1996: JP-080702.
PA (WEIP) MEIJI MILK PROD CO LTD.
PA Dairiki K, Iwama A, Kino K, Kume A, Some T.
DR WPI: 97-470495/43.
PT Peptide immunotherapeutic agent to treat allergic diseases - contains multi-epitope peptide containing T cell epitope regions from different allergens.
PS Claim 9: Page 32: 58pp: Japanese.
CC The present sequence represents a multi-epitope peptide which is used as a new immunotherapeutic agent. It comprises T cell epitope regions from 2 or more different allergens (preferably linked via arginine or lysine dimers), where the T cell epitope regions have a positivity index greater than 100 as measured in a patient group responding to the allergen; have at least 70% reactivity with lymphocytes from patients responding to the allergen; and are not reactive with immunoglobulin E (IgE) antibodies from patients responsive to the allergen. The agent can be used to prevent and treat a wide variety of allergic diseases, e.g. by desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
SQ Sequence 31 AA:
Query Match: 100.0%; Score 64; DB 1; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.25e+00;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DI 23 FIKRVSNVI 31
:|||||
QY 1 FIKRVSNVI 9

RESULT 9
ID W27369 standard: peptide: AA.
AC W27369
DI 24-MAR-1998 (first entry)
DE Multi-epitope peptide used as immunotherapeutic agent #1.
KW Multi-epitope peptide; immunotherapeutic agent; allergic disease.
KW T-cell epitope region; allergen; lymphocyte; immunoglobulin E.
OS Synthetic.
PN W09732600-A1.
PC 12-SEP-1997.
PR 10-MAR-1996: JP-0807-2.
PA (WEIP) MEIJI MILK PROD CO LTD.
PA Dairiki K, Iwama A, Kino K, Kume A, Some T.
DR WPI: 97-470495/43.
PT Peptide immunotherapeutic agent to treat allergic diseases - contains multi-epitope peptide containing T cell epitope regions from different allergens.
PS Claim 6: Page 31: 58pp: Japanese.
CC The present sequence represents a multi-epitope peptide which is used as a new immunotherapeutic agent. It comprises T cell epitope regions from 2 or more different allergens (preferably linked via arginine or lysine dimers), where the T cell epitope regions have a positivity index greater than 100 as measured in a patient group responding to the allergen; have at least 70% reactivity with lymphocytes from patients responding to the allergen; and are not reactive with immunoglobulin E (IgE) antibodies from patients responsive to the allergen. The agent can be used to prevent and treat a wide variety of allergic diseases, e.g. by desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
SQ Sequence 80 AA:
Query Match: 100.0%; Score 64; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 1.25e+00;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DI 17 FIKRVSNVI 25

```

QY      1 FIKRVSNVI 9
      |||||
RESULT  10
ID W27370 standard; peptide: 105 AA.
AC W27370
DE 24-MAR-1998 (first entry)
KW Multi-epitope peptide used as immunotherapeutic agent #2.
KW Multi-epitope peptide: immunotherapeutic agent; allergic disease.
KW T-cell epitope region: allergen; lymphocyte; immunoglobulin E.
OS Synthetic.
PN W09732600-A1.
PD 12-SEP-1997.
PR 10-MAR-1997: J00740.
PR 10-MAR-1996: JP-080702.
PA (MEIP ) MEIJI MILK PROD CO LTD.
PA Dairiki K, Iwama A, Kino K, Kume A, Sone T;
PA WPI: 97-470495/43.
PT Peptide immunotherapeutic agent to treat allergic diseases -
PT contains multi-epitope peptide containing T cell epitope regions
PT from different allergens
PS Claim 6: Page 31: 58pp; Japanese.
CC The present sequence represents a multi-epitope peptide which is used as
CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2
CC or more different allergens (preferably linked via arginine or lysine
CC dimers), where the T cell epitope regions have a positivity index
CC greater than 100 as measured in a patient group responding to the
CC allergen; have at least 70% reactivity with lymphocytes from patients
CC responding to the allergen; and are not reactive with immunoglobulin E
CC (IgE) antibodies from patients responsive to the allergen. The agent can
CC be used to prevent and treat a wide variety of allergic diseases, e.g. by
CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
SQ Sequence 105 AA.

Query Match 100.0% Score 64: DB 1: Length 105;
Best Local Similarity 100.0% Pred. No. 1.25e-00;
Matches 9: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY      1 FIKRVSNVI 9
      |||||
RESULT  11
ID W27371 standard; peptide: 114 AA.
AC W27371
DE 24-MAR-1998 (first entry)
KW Multi-epitope peptide used as immunotherapeutic agent #3.
KW Multi-epitope peptide: immunotherapeutic agent; allergic disease.
KW T-cell epitope region: allergen; lymphocyte; immunoglobulin E.
OS Synthetic.
PN W09732600-A1.
PD 12-SEP-1997.
PR 10-MAR-1997: J00740.
PR 10-MAR-1996: JP-080702.
PA (MEIP ) MEIJI MILK PROD CO LTD.
PA Dairiki K, Iwama A, Kino K, Kume A, Sone T;
PA WPI: 97-470495/43.
PT Peptide immunotherapeutic agent to treat allergic diseases -
PT contains multi-epitope peptide containing T cell epitope regions
PT from different allergens
PS Claim 6: Page 32: 58pp; Japanese.
CC The present sequence represents a multi-epitope peptide which is used as
CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2
CC or more different allergens (preferably linked via arginine or lysine
CC dimers), where the T cell epitope regions have a positivity index
CC greater than 100 as measured in a patient group responding to the
CC allergen; have at least 70% reactivity with lymphocytes from patients
CC responding to the allergen; and are not reactive with immunoglobulin E
CC (IgE) antibodies from patients responsive to the allergen. The agent can
CC be used to prevent and treat a wide variety of allergic diseases, e.g. by
CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
SQ Sequence 105 AA.

```

```

SQ      Sequence 134 AA:
Query Match 100.0% Score 64: DB 1: Length 134;
Best Local Similarity 100.0% Pred. No. 1.25e-00;
Matches 9: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

DB      17 FIKRVSNVI 25
      |||||
QY      1 FIKRVSNVI 9
      |||||
RESULT  12
ID R75388 standard; protein: 353 AA.
AC R75388;
DE 12-MAR-1996 (first entry)
DE Japanese cedar pollen allergen Cryj 1.
KW Japanese cedar; pollen allergen; Cryj 1; T-cell epitope; peptides.
KW Prevention; treatment; cryptomeria pollinosis.
OS Cryptomeria japonica.
PH Key: Location/Qualifiers
FT peptide 61..75
FT /note= "T-cell epitope peptide"
FT peptide 91..105
FT /note= "T-cell epitope peptide"
FT peptide 106..120
FT /note= "T-cell epitope peptide"
FT peptide 146..160
FT /note= "T-cell epitope peptide"
FT peptide 211..225
FT /note= "T-cell epitope peptide"
FT peptide 326..340
FT /note= "T-cell epitope peptide"
FT peptide 335..346
FT /note= "T-cell epitope peptide"
PN J07118295-A.
PD 09-MAY-1995.
PR 20-OCT-1993: 262626.
PR 20-OCT-1993: JP-262626.
PA (MEIP ) MEIJI MILK PROD CO LTD.
PA WPI: 95-203834/27.
PT New cryptomeria pollen allergen T-cell epitope peptide - used for
PT prevention, treatment and investigation of Japanese cedar pollinosis
PT Disclosure; Figs 1-2: 8pp; Japanese.
CC R75388 is the Japanese cedar pollen allergen Cryj 1, from which the
CC T-cell epitope peptides R89249-R89295 were derived. The peptides
CC can be used for the prevention and treatment of cryptomeria
CC pollinosis, and also for the investigation of pollinosis.
SQ Sequence 353 AA:

Query Match 100.0% Score 64: DB 1: Length 453;
Best Local Similarity 100.0% Pred. No. 1.25e-00;
Matches 9: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

DB      109 FIKRVSNVI 117
      |||||
QY      1 FIKRVSNVI 9
      |||||
RESULT  13
ID R81587 standard; protein: 353 AA.
AC R81587;
DE 24-MAY-1996 (first entry)
DE Cedar pollen allergen B.
KW Cedar; pollen; allergen; immunoglobulin E; IgE; T-cell epitope;
KW antibody; pollinosis; therapy; immunotherapy.
OS Cryptomeria japonica.
PN EP-700929-A2.
PD 13-MAR-1996.
PR 08-SEP-1995: 306295.
PR 10-SEP-1994: JP-242137.
PR 14-JUL-1995: JP-200221.
PR 14-JUL-1995: JP-200204.
PA (HAYB ) HAYASHIARA SEIBUTSU KAGAKU.

```

Page 5

CC TOPOLOGY: LINEAR
CC MOLECULE TYPE: PROTEIN
SQ SEQUENCE 317 AA: 37355 MW: 525334 CN:

Query Match 78.1% Score 50: DB 3: Length 317:
Best Local Similarity 66.7% Pred. No. 4.45e-01:
Matches 6: Conservative 2: Mismatches 1: Indels 0: Gaps 0:

DB 72 FPKRVDEVI 80

QY 1 FPKRVSNVI 9

RESULT 2
ID US-08-951-889-2 STANDARD: PRT: 317 AA:

XX AC xxxxxx

XX DT :|||:

XX XX

DE DE

XX SEQUENCE 2: Application US/08951889

CC SEQUENCE 2: Application US/08951889

CC PATENT NO. 6008332

CC GENERAL INFORMATION:

CC APPLICANT: Mathur, E. et al.

CC TITLE OF INVENTION: Carboxymethyl Cellulase from

CC TITLE OF INVENTION: Theriotoga Maritima

CC NUMBER OF SEQUENCES: 4

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,

CC ADDRESSEE: CECCHI, STEWART & OLSTEIN

CC CITY: ROSELAND

CC STATE: NEW JERSEY

CC COUNTRY: USA

CC ZIP: 07068

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: 3.5 INCH DISKETTE

CC COMPUTER: IBM PS/2

CC OPERATING SYSTEM: MS-DOS

CC SOFTWARE: WORD PERFECT 5.1

CC CURRENT APPLICATION DATA:

CC FILING DATE: August 23, 1995

CC CLASSIFICATION: 4.35

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US/08/951-889

CC FILING DATE: August 23, 1995

CC NAME: FERRARO, GREGORY D.

CC REGISTRATION NUMBER: 36,134

CC REFERENCE/DOCKET NUMBER: 331400-20

CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 201-994-1700
CC TELEFAX: 201-994-1744
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 317 AMINO ACIDS
CC TYPE: AMINO ACID
CC STRANDEDNESS:
CC TOPOLOGY: LINEAR
CC MOLECULE TYPE: PROTEIN
SQ SEQUENCE 317 AA: 37355 MW: 525334 CN:

Query Match 78.1% Score 50: DB 3: Length 317:
Best Local Similarity 66.7% Pred. No. 4.45e-01:
Matches 6: Conservative 2: Mismatches 1: Indels 0: Gaps 0:

DB 72 FPKRVDEVI 80

QY 1 FPKRVSNVI 9

RESULT 2
ID US-08-951-889-2 STANDARD: PRT: 317 AA:

XX AC xxxxxx

XX DT :|||:

XX XX

DE DE

XX SEQUENCE 2: Application US/08951889

CC SEQUENCE 2: Application US/08951889

CC PATENT NO. 6008332

CC GENERAL INFORMATION:

CC APPLICANT: Mathur, E. et al.

CC TITLE OF INVENTION: Carboxymethyl Cellulase from

CC TITLE OF INVENTION: Theriotoga Maritima

CC NUMBER OF SEQUENCES: 4

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,

CC ADDRESSEE: CECCHI, STEWART & OLSTEIN

CC CITY: ROSELAND

CC STATE: NEW JERSEY

CC COUNTRY: USA

CC ZIP: 07068

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: 3.5 INCH DISKETTE

CC COMPUTER: IBM PS/2

CC OPERATING SYSTEM: MS-DOS

CC SOFTWARE: WORD PERFECT 5.1

CC CURRENT APPLICATION DATA:

CC FILING DATE: August 23, 1995

CC CLASSIFICATION: 4.35

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US/08/951-889

CC FILING DATE: August 23, 1995

CC NAME: FERRARO, GREGORY D.

CC REGISTRATION NUMBER: 36,134

CC REFERENCE/DOCKET NUMBER: 331400-20

RESULT 3
ID US-09-066-075-2 STANDARD: PRT: 317 AA:

XX AC xxxxxx

XX DT :|||:

XX XX

DE DE

XX SEQUENCE 2: Application US/09046075

CC SEQUENCE 2: Application US/09046075

CC PATENT NO. 5925749

CC GENERAL INFORMATION:

CC APPLICANT: Mathur, E. et al.

CC TITLE OF INVENTION: Carboxymethyl Cellulase from

CC TITLE OF INVENTION: Theriotoga Maritima

CC NUMBER OF SEQUENCES: 4

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,

CC ADDRESSEE: CECCHI, STEWART & OLSTEIN

CC CITY: ROSELAND

CC STATE: NEW JERSEY

CC COUNTRY: USA

CC ZIP: 07068

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: 3.5 INCH DISKETTE

CC COMPUTER: IBM PS/2

CC OPERATING SYSTEM: MS-DOS

CC SOFTWARE: WORD PERFECT 5.1

CC CURRENT APPLICATION DATA:

CC FILING DATE: August 23, 1995

CC CLASSIFICATION:

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US/08/9518-615

CC FILING DATE: August 23, 1995

CC NAME: FERRARO, GREGORY D.

CC REGISTRATION NUMBER: 36,134

CC REFERENCE/DOCKET NUMBER: 331400-20

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: 201-994-1700

CC TELEFAX: 201-994-1744

CC INFORMATION FOR SEQ ID NO: 2:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 317 AMINO ACIDS

CC TYPE: AMINO ACID
CC STRANDEDNESS:
CC TOPOLOGY: LINEAR
CC MOLECULE TYPE: PROTEIN
SQ SEQUENCE 317 AA: 37355 MW: 525334 CN:

Query Match 78.1% Score 50: DB 2: Length 317:
Best Local Similarity 66.7% Pred. No. 4.45e-01:
Matches 6: Conservative 2: Mismatches 1: Indels 0: Gaps 0:

DB 72 FPKRVDEVI 80

QY 1 FPKRVSNVI 9

RESULT 4
ID US-08-459-263-4 STANDARD: PRT: 255 AA:

XX AC xxxxxx

XX DT :|||:

XX XX

DE DE

XX SEQUENCE 4: Application US/08459263

CC SEQUENCE 4: Application US/08459263

CC PATENT NO. 5792642

CC GENERAL INFORMATION:

CC APPLICANT: Mathur, E. et al.

CC TITLE OF INVENTION: Carboxymethyl Cellulase from

CC TITLE OF INVENTION: Theriotoga Maritima

CC NUMBER OF SEQUENCES: 4

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,

CC ADDRESSEE: CECCHI, STEWART & OLSTEIN

CC CITY: ROSELAND

CC STATE: NEW JERSEY

CC COUNTRY: USA

CC ZIP: 07068

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: 3.5 INCH DISKETTE

CC COMPUTER: IBM PS/2

CC OPERATING SYSTEM: MS-DOS

CC SOFTWARE: WORD PERFECT 5.1

CC CURRENT APPLICATION DATA:

CC FILING DATE: August 23, 1995

CC CLASSIFICATION:

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US/08/9518-615

CC FILING DATE: August 23, 1995

CC NAME: FERRARO, GREGORY D.

CC REGISTRATION NUMBER: 36,134

CC REFERENCE/DOCKET NUMBER: 331400-20

CC APPLICANT: RASMUSSEN, Beth A
CC APPLICANT: TALLY, Francis P
CC TITLE OF INVENTION: DNA SEQUENCES AND AMINO ACID SEQUENCES
CC TITLE OF INVENTION: OF CLASS B BETA-LACTAMASE ENZYMES FROM BACTERIOIDES
CC TITLE OF INVENTION: FRAGILIS
CC NUMBER OF SEQUENCES: 9
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Darby & Darby PC
CC STREET: 805 Third Avenue
CC CITY: New York
CC STATE: NY
CC COUNTRY: US
CC ZIP: 10022
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/459,263
CC FILING DATE: 02-JUN-1995
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Robinson, Joseph R
CC REGISTRATION NUMBER: 33,448
CC REFERENCE/DOCKET NUMBER: 0646/18026-US2
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 212-527-7700
CC TELEFAX: 212-753-6237
CC TELEX: 236687
CC INFORMATION FOR SEQ ID NO: 4:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 255 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC ORIGINAL SOURCE:
CC ORGANISM: Bacillus cereus
CC IMMEDIATE SOURCE:
CC CLONE: B-LACTAMASE
CC SEQUENCE 255 AA: 27847 MW: 349604 CN:

Query Match 76.6% Score 49: DB 1: Length 255:
Best Local Similarity 66.7% Pred. No. 5.76e-01:
Matches 6: Conservative 2: Mismatches 1: Indels 0: Gaps 0:

DB 103 FKKRVTDVI 111
1111111111
QY 1 FKKRVSNVI 9

RESULT 5
ID US-08-459-264-4 STANDARD: PRT: 255 AA.
XX
AC xxxxxx
XX
DT
XX
DE
XX
Sequence 4, Application US/08459264
Patent No. 5703340
GENERAL INFORMATION:
CC APPLICANT: RASMUSSEN, Beth A
CC APPLICANT: TALLY, Francis P
CC APPLICANT: GLUZMAN, Yakov
CC TITLE OF INVENTION: DNA SEQUENCES AND AMINO ACID SEQUENCES
CC TITLE OF INVENTION: OF CLASS B BETA-LACTAMASE ENZYMES FROM BACTERIOIDES
CC NUMBER OF SEQUENCES: 9
CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Darby & Darby PC
CC STREET: 805 Third Avenue
CC CITY: New York
CC STATE: NY
CC COUNTRY: US
CC ZIP: 10022
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/459,264
CC FILING DATE: 02-JUN-1995
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Robinson, Joseph R
CC REGISTRATION NUMBER: 33,448
CC REFERENCE/DOCKET NUMBER: 0646/18026-US3
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 212-527-7700
CC TELEFAX: 212-753-6237
CC TELEX: 236687
CC INFORMATION FOR SEQ ID NO: 4:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 255 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC ORIGINAL SOURCE:
CC ORGANISM: Bacillus cereus
CC IMMEDIATE SOURCE:
CC CLONE: B-LACTAMASE
CC SEQUENCE 255 AA: 27847 MW: 349604 CN:

Query Match 76.6% Score 49: DB 1: Length 255:
Best Local Similarity 66.7% Pred. No. 5.76e-01:
Matches 6: Conservative 2: Mismatches 1: Indels 0: Gaps 0:

DB 103 FKKRVTDVI 111
1111111111
QY 1 FKKRVSNVI 9

RESULT 6
ID US-08-978-182-1 STANDARD: PRT: 424 AA.
XX
AC xxxxxx
XX
DT
XX
DE
XX
Sequence 1, Application US/08978182
Patent No. 5849556
GENERAL INFORMATION:
CC APPLICANT: Hillman, Jennifer L.
CC APPLICANT: Yue, Henry
CC APPLICANT: Guegler, Karl J.
CC APPLICANT: Kaser, Matthew
CC APPLICANT: Mathur, Preete
CC TITLE OF INVENTION: HUMAN GROWTH-RELATED CDC10 HOMOLOG
CC NUMBER OF SEQUENCES: 5
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Incyte Pharmaceuticals, Inc.
CC STREET: 3174 Porter Drive
CC CITY: Palo Alto
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94304
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette

CC COMPUTER: IBM Compatible
 CC OPERATING SYSTEM: DOS
 CC SOFTWARE: FastSeq for Windows Version 2.0
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/978,182
 CC FILING DATE: Herewith
 CC CLASSIFICATION: 435
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER:
 CC FILING DATE:
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Billings, Lucy J.
 CC REGISTRATION NUMBER: 36,749
 CC REFERENCE/DOCKET NUMBER: PF-0426 US
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: 650-855-0555
 CC TELEFAX: 650-845-4156
 CC TELEX:
 CC INFORMATION FOR SEQ ID NO: 1:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 424 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS: Single
 CC TOPOLOGY: linear
 CC IMMEDIATE SOURCE:
 CC LIBRARY: TLYMN0106
 CC CLONE: 3003826
 CC SEQUENCE 424 AA: 47734 MW: 90,865.0N:
 CC
 CC Query Match 73.4% Score 471 DB-2 Length 424:
 CC Best Local Similarity 55.6% Pred. No. 9,576-01:
 CC Matches 5: Conservative 4: Mismatches 0: Indels 0: Gaps 0:
 CC
 CC DB 265 EMKRSKV 273
 CC 1111111
 CC QY 1 FIKVSNVI 9
 CC
 CC RESULT 7
 CC ID US-09-205-487-1 STANDARD: PFI: 424 AA.
 CC AC xxxxxx
 CC UT
 CC SE Sequence 1 Application 18/092,568:
 CC Sequence 1 Application 18/092,568:
 CC Patent No. 5,452,214
 CC GENERAL INFORMATION:
 CC APPLICANT: Billman, Jonathan D.
 CC APPLICANT: Y. P. Henry
 CC APPLICANT: Greider, Karl J.
 CC APPLICANT: Kaser, Matthew
 CC APPLICANT: Mathur, Brooke
 CC TITLE OF INVENTION: HUMAN GROWTH-RELATED CD-10 HOMOLOG
 CC NUMBER OF SEQUENCES: 1
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Cyclope Pharmaceuticals, Inc.
 CC STREET: 3174 Porter Drive
 CC CITY: Palo Alto
 CC STATE: CA
 CC COUNTRY: USA
 CC ZIP: 94304
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Diskette
 CC COMPUTER: IBM Compatible
 CC OPERATING SYSTEM: DOS
 CC SOFTWARE: FastSeq for Windows Version 2.0
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/09/205,481
 CC FILING DATE:
 CC CLASSIFICATION:

CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/978,182
 CC FILING DATE:
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Billings, Lucy J.
 CC REGISTRATION NUMBER: 36,749
 CC REFERENCE/DOCKET NUMBER: PF-0426 US
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: 650-855-0555
 CC TELEFAX: 650-845-4156
 CC TELEX:
 CC INFORMATION FOR SEQ ID NO: 1:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 424 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS: Single
 CC TOPOLOGY: linear
 CC IMMEDIATE SOURCE:
 CC LIBRARY: TLYMN0106
 CC CLONE: 3003826
 CC SEQUENCE 424 AA: 47734 MW: 90,865.0N:
 CC
 CC Query Match 73.4% Score 471 DB-2 Length 424:
 CC Best Local Similarity 55.6% Pred. No. 9,576-01:
 CC Matches 5: Conservative 4: Mismatches 0: Indels 0: Gaps 0:
 CC
 CC DB 265 EMKRSKV 273
 CC 1111111
 CC QY 1 FIKVSNVI 9
 CC
 CC RESULT 8
 CC ID FCI-US95-10398-80 STANDARD: PFI: 192 AA.
 CC AC xxxxxx
 CC UT
 CC SE Sequence 80, Application 90/ 4510-48
 CC Sequence 80, Application 90/ 4510-48
 CC GENERAL INFORMATION:
 CC APPLICANT: BURKH, J. MICHAEL P. H. AND
 CC APPLICANT: PURCELL, R.J.
 CC TITLE OF INVENTION: NOVEL HHH AND DERIVED
 CC TITLE OF INVENTION: AMIN ACID SEQUENCES OF THE ENVELOPE 1 AND
 CC TITLE OF INVENTION: CHAIN LAIN PROTEIN LIPID-PROTEIN COMPLEXES
 CC TITLE OF INVENTION: AN ANTI-HERPES VIRUS RECEPTOR
 CC TITLE OF INVENTION: AN ANTI-HERPES VIRUS RECEPTOR
 CC NUMBER OF SEQUENCES: 1
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: MORGAN & FINNIGAN
 CC STREET: 345 PARK AVENUE
 CC CITY: NEW YORK
 CC STATE: NEW YORK
 CC COUNTRY: USA
 CC ZIP: 10154
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: FLOPPY DISK
 CC COMPUTER: IBM PC COMPATIBLE
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: WORDPERFECT 5.1
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: 95/US95/10398
 CC FILING DATE: 15-AUG-1995
 CC CLASSIFICATION:
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 95/085,428
 CC FILING DATE: 29 JUNE 1993
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 95/290/655
 CC FILING DATE: 15 AUGUST 1994
 CC ATTORNEY/AGENT INFORMATION:

SEQUENCE CHARACTERISTICS:

145 vMkVP:V+ 153

Jot time : 5 secs

QY 1 FIKRVSNVI 9

RESULT 6
ID US-08-350-225-36 STANDARD: PPT: 20 AA.

XX AC xxxxxx

XX DT

XX DE

XX Sequence 36, Application US/08350225

XX Sequence 36, Application US/08350225

XX GENERAL INFORMATION:

CC APPLICANT: Griffith, Irwin J.

CC APPLICANT: Pollock, Joanne,

CC APPLICANT: Bond, Julian F.

CC APPLICANT: Garman, Richard D.

CC APPLICANT: Kuo, Mei-Chang.

CC APPLICANT: Yeung, Siu-mei H.

CC APPLICANT: Brauer, Andrew.

CC APPLICANT: Exley, Mark A.

CC APPLICANT: Powers, Steven P.

CC APPLICANT: Griffith, Irwin J.

CC TITLE OF INVENTION: Allergenic Proteins And Peptides From

CC TITLE OF INVENTION: Japanese Cedar Pollen

CC NUMBER OF SEQUENCES: 261

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.

CC STREET: 610 Lincoln St

CC CITY: Waltham

CC STATE: MA

CC COUNTRY: USA

CC ZIP: 02154

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: Patent Release #1.0, Version #1.25

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08350225

CC FILING DATE: December 6, 1994

CC CLASSIFICATION: 424

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: 08/226,248

CC FILING DATE: April 8, 1994

CC APPLICATION NUMBER: 07/938,990

CC FILING DATE: September 1, 1992

CC APPLICATION NUMBER: 08/05933/0139

CC FILING DATE: January 15, 1995

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Darlene A. Vastore

CC REGISTRATION NUMBER: 35,729

CC REFERENCE/DOCKET NUMBER: 025 6 US (IM1-028CP2)

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: (617) 466-6000

CC TELEFAX: (617) 466-6040

CC INFORMATION FOR SEQ ID NO: 36:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 20 amino acids

CC TYPE: amino acid

CC TOPOLOGY: linear

CC MOLECULE TYPE: peptide

CC FRAGMENT TYPE: internal

CC SEQUENCE 20 AA: 2080 MW: 2304 CN:

Query Match 100.0% Score 64: DB 8: Length 20:

Best Local Similarity 100.0% Pred. No. 2 93e-00:

Matches 9: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

DB 9 FIKRVSNVI 17

QY 1 FIKRVSNVI 9

RESULT 7

ID US-09-467-597-36 STANDARD: PPT: 20 AA.

XX AC xxxxxx

XX DT

XX DE

XX Sequence 36, Application US/08467597

XX Sequence 36, Application US/08467597

XX GENERAL INFORMATION:

CC APPLICANT: Griffith, Irwin J.

CC APPLICANT: Pollock, Joanne,

CC APPLICANT: Bond, Julian F.

CC APPLICANT: Garman, Richard D.

CC APPLICANT: Kuo, Mei-Chang.

CC APPLICANT: Yeung, Siu-mei H.

CC APPLICANT: Brauer, Andrew.

CC APPLICANT: Exley, Mark A.

CC APPLICANT: Powers, Steven P.

CC APPLICANT: Griffith, Irwin J.

CC TITLE OF INVENTION: Allergenic Proteins And Peptides From

CC TITLE OF INVENTION: Japanese Cedar Pollen

CC NUMBER OF SEQUENCES: 261

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.

CC STREET: 610 Lincoln St

CC CITY: Waltham

CC STATE: MA

CC COUNTRY: USA

CC ZIP: 02154

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-D /MS-DOS

CC SOFTWARE: Patent Release #1.0, Version #1.25

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08467597

CC FILING DATE: June 6, 1995

CC CLASSIFICATION: 424

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: 08/350,225

CC FILING DATE: December 6, 1994

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Darlene A. Vastore

CC REGISTRATION NUMBER: 35,729

CC REFERENCE/DOCKET NUMBER: 025 6 US (IM1-28CP2)

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: (617) 227-7400

CC TELEFAX: (617) 227-7441

CC INFORMATION FOR SEQ ID NO: 36:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 20 amino acids

CC TYPE: amino acid

CC TOPOLOGY: linear

CC MOLECULE TYPE: peptide

CC FRAGMENT TYPE: internal

CC SEQUENCE 20 AA: 2080 MW: 2304 CN:

Query Match 100.0% Score 64: DB 10: Length 20:

Best Local Similarity 100.0% Pred. No. 2 93e-00:

Matches 9: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

DB 9 FIKRVSNVI 17

QY 1 FIKRVSNVI 9

RESULT 8

ID US-09-142-524-4 STANDARD: PPT: 31 AA.

XX AC xxxxxx


```

RESULT 11
ID US-08-468-940-62 STANDARD: PRT: 60 AA.
XX
AC
XX
XX
DT
XX
DE Sequence 52, Application: US/08468940
XX
CC Sequence 52, Application US/08468940
CC GENERAL INFORMATION:
CC APPLICANT: Griffith, Irwin J.
CC APPLICANT: Pollock, Joanne;
CC APPLICANT: Bond, Julian F.;
CC APPLICANT: Garman, Richard D.;
CC APPLICANT: Kuo, Mei-Chang;
CC APPLICANT: Yeung, Siu-mei H.;
CC APPLICANT: Brauer, Andrew;
CC APPLICANT: Exley, Mark A.;
CC APPLICANT: Powers, Steven P.;
CC TITLE OF INVENTION: Allergic Proteins And Peptides From
CC TITLE OF INVENTION: Japanese Cedar Pollen
CC NUMBER OF SEQUENCES: 251
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
CC STREET: 610 Lincoln St
CC CITY: Waltham
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02154
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08468-940
CC FILING DATE:
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/350,225
CC FILING DATE: December 6, 1994
CC APPLICATION NUMBER: 08/226,248
CC FILING DATE: April 8, 1994
CC APPLICATION NUMBER: 07/538,940
CC FILING DATE: September 1, 1992
CC APPLICATION NUMBER: PCT/US93/08134
CC FILING DATE: January 15, 1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Darlene A. Vansluis
CC REGISTRATION NUMBER: 35,729
CC REFERENCE/DOCKET NUMBER: 025.6 US (IM-028CP2)
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 466-6000
CC TELEFAX: (617) 466-6040
CC INFORMATION FOR SEQ ID NO: 62:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 60 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC FRAGMENT TYPE: internal
CC SEQUENCE 60 AA: 6644 MW: 19464 CN:
SQ
Query Match 100.0% Score 64: DB 10: Length 60:
Best Local Similarity 100.0% Pred. No. 2,93e+00:
Matches 9: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
Db 49 FIKRVSNI 57
QY : FIKRVSNI 9

```

```

RESULT 12
ID US-08-226-248A-62 STANDARD: PRT: 60 AA.
XX
AC
XX
XX
DT
XX
DE Sequence 62, Application: US/08226248A
XX
CC Sequence 62, Application US/08226248A
CC GENERAL INFORMATION:
CC APPLICANT: Griffith, Irwin J.
CC APPLICANT: Pollock, Joanne;
CC APPLICANT: Bond, Julian F.;
CC APPLICANT: Garman, Richard D.;
CC APPLICANT: Kuo, Mei-Chang;
CC APPLICANT: Yeung, Siu-mei H.;
CC APPLICANT: Brauer, Andrew;
CC APPLICANT: Exley, Mark A.;
CC APPLICANT: Powers, Steven P.;
CC TITLE OF INVENTION: Allergic Proteins And Peptides From
CC TITLE OF INVENTION: Japanese Cedar Pollen
CC NUMBER OF SEQUENCES: 201
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
CC STREET: 610 Lincoln St
CC CITY: Waltham
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02154
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/226.248A
CC FILING DATE: April 8, 1994
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/538,990
CC FILING DATE: September 1, 1992
CC APPLICATION NUMBER: PCT/US93/08146
CC FILING DATE: January 15, 1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Darlene A. Vansluis
CC REGISTRATION NUMBER: 35,729
CC REFERENCE/DOCKET NUMBER: 025.6 US (IM-028CP)
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 466-6000
CC TELEFAX: (617) 466-6040
CC INFORMATION FOR SEQ ID NO: 62:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 60 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC FRAGMENT TYPE: internal
CC SEQUENCE 60 AA: 6644 MW: 19464 CN:
SQ
Query Match 100.0% Score 64: DB 7: Length 60:
Best Local Similarity 100.0% Pred. No. 2,93e+00:
Matches 9: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
Db 49 FIKRVSNI 57
QY : FIKRVSNI 9
RESULT 13
ID US-08-467-697-62 STANDARD: PRT: 60 AA.

```

```

XX      xxxxxx
AC
XX
DT
XX      Sequence 62, Application US/08467697
XX      Sequence 62, Application US/08467697
CC      GENERAL INFORMATION:
CC      APPLICANT: Griffith, Irwin J.
CC      APPLICANT: Pollock, Joanne
CC      APPLICANT: Bond, Julian F.
CC      APPLICANT: Garmar, Richard D.
CC      APPLICANT: Kuo, Mei-Chang
CC      APPLICANT: Yeung, Siu-mei H.
CC      APPLICANT: Brauer, Andrew
CC      APPLICANT: Exley, Mark A.
CC      APPLICANT: Powers, Steven P.
CC      TITLE OF INVENTION: Allergenic Proteins And Peptides From
CC      TITLE OF INVENTION: Japanese Cedar Pollen
CC      NUMBER OF SEQUENCES: 261
CC      CORRESPONDENCE ADDRESS:
CC      ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
CC      STREET: 610 Lincoln St.
CC      CITY: Waltham
CC      STATE: MA
CC      COUNTRY: USA
CC      ZIP: 02154
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: Patent In Release #1.0, Version #1.25
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: US/08/467,697
CC      FILING DATE: June 6, 1995
CC      CLASSIFICATION: 424
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: 08/350,225
CC      FILING DATE: December 6, 1994
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: Jane E. Remillard
CC      REGISTRATION NUMBER: 38,872
CC      REFERENCE/DOCKET NUMBER: 025.6 USD4 (IM-028CPD4)
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: (617) 227-7400
CC      TELEFAX: (617) 227-5941
CC      INFORMATION FOR SEQ ID NO: 62:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 60 amino acids
CC      MOLECULE TYPE: linear
CC      FRAGMENT TYPE: internal
CC      SEQUENCE 60 AA: 6644 MW: 19464 CN:
SQ
      Query Match 100.0%; Score 64; DB 10; Length 60;
      Best Local Similarity 100.0%; Pred. No. 2,93e-00;
      Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB      43 FIKRVSNVI 57
QY      1 FIKRVSNVI 9
      RESULT 14
      ID US-08-467-023-62 STANDARD; PRT: 60 AA.
      AC xxxxxx
      XX
      DT
      DE
      Sequence 62, Application US/08467697
      GENERAL INFORMATION:
      APPLICANT: Griffith, Irwin J.
      APPLICANT: Pollock, Joanne
      APPLICANT: Bond, Julian F.
      APPLICANT: Garmar, Richard D.
      APPLICANT: Kuo, Mei-Chang
      APPLICANT: Yeung, Siu-mei H.
      APPLICANT: Brauer, Andrew
      APPLICANT: Exley, Mark A.
      APPLICANT: Powers, Steven P.
      TITLE OF INVENTION: Allergenic Proteins And Peptides From
      TITLE OF INVENTION: Japanese Cedar Pollen
      NUMBER OF SEQUENCES: 261
      CORRESPONDENCE ADDRESS:
      ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
      STREET: 610 Lincoln St.
      CITY: Waltham
      STATE: MA
      COUNTRY: USA
      ZIP: 02154
      COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: Patent In Release #1.0, Version #1.25
      CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/08/467,697
      FILING DATE: June 6, 1995
      CLASSIFICATION: 424
      PRIOR APPLICATION DATA:
      APPLICATION NUMBER: 08/350,225
      FILING DATE: December 6, 1994
      ATTORNEY/AGENT INFORMATION:
      NAME: Jane E. Remillard
      REGISTRATION NUMBER: 38,872
      REFERENCE/DOCKET NUMBER: 025.6 USD4 (IM-028CPD4)
      TELECOMMUNICATION INFORMATION:
      TELEPHONE: (617) 227-7400
      TELEFAX: (617) 227-5941
      INFORMATION FOR SEQ ID NO: 62:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 60 amino acids
      MOLECULE TYPE: linear
      FRAGMENT TYPE: internal
      SEQUENCE 60 AA: 6644 MW: 19464 CN:
SQ
      Query Match 100.0%; Score 64; DB 10; Length 60;
      Best Local Similarity 100.0%; Pred. No. 2,93e-00;
      Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB      43 FIKRVSNVI 57
QY      1 FIKRVSNVI 9
      RESULT 14
      ID US-08-467-023-62 STANDARD; PRT: 60 AA.
      AC xxxxxx
      XX
      DT
      DE
      Sequence 62, Application US/08467697
      GENERAL INFORMATION:
      APPLICANT: Griffith, Irwin J.
      APPLICANT: Pollock, Joanne
      APPLICANT: Bond, Julian F.

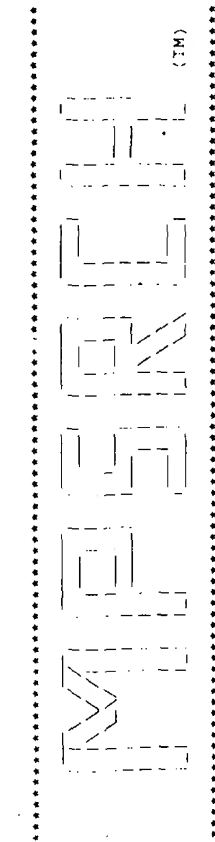
```

```

XX      Sequence 62, Application US/ 467023
XX      GENERAL INFORMATION:
XX      APPLICANT: Griffith, Irwin J.
XX      APPLICANT: Pollock, Joanne
XX      APPLICANT: Bond, Julian F.
XX      APPLICANT: Garmar, Richard D.
XX      APPLICANT: Kuo, Mei-Chang
XX      APPLICANT: Yeung, Siu-mei H.
XX      APPLICANT: Brauer, Andrew
XX      APPLICANT: Exley, Mark A.
XX      APPLICANT: Powers, Steven P.
XX      TITLE OF INVENTION: Allergenic Proteins And Peptides From
XX      TITLE OF INVENTION: Japanese Cedar Pollen
XX      NUMBER OF SEQUENCES: 261
XX      CORRESPONDENCE ADDRESS:
XX      ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
XX      STREET: 610 Lincoln St.
XX      CITY: Waltham
XX      STATE: MA
XX      COUNTRY: USA
XX      ZIP: 02154
XX      COMPUTER READABLE FORM:
XX      MEDIUM TYPE: Floppy disk
XX      COMPUTER: IBM PC compatible
XX      OPERATING SYSTEM: PC-DOS/MS-DOS
XX      SOFTWARE: Patent In Release #1.0, Version #1.25
XX      CURRENT APPLICATION DATA:
XX      APPLICATION NUMBER: US/08/467,023
XX      FILING DATE: June 6, 1995
XX      CLASSIFICATION: 424
XX      PRIOR APPLICATION DATA:
XX      APPLICATION NUMBER: 08/350,225
XX      FILING DATE: December 6, 1994
XX      ATTORNEY/AGENT INFORMATION:
XX      NAME: Jane E. Remillard
XX      REGISTRATION NUMBER: 38,872
XX      REFERENCE/DOCKET NUMBER: 025.6 USD2 (IM-028CPD2)
XX      TELECOMMUNICATION INFORMATION:
XX      TELEPHONE: (617) 227-7400
XX      TELEFAX: (617) 227-5941
XX      INFORMATION FOR SEQ ID NO: 62:
XX      SEQUENCE CHARACTERISTICS:
XX      LENGTH: 60 amino acids
XX      MOLECULE TYPE: linear
XX      FRAGMENT TYPE: internal
XX      SEQUENCE 60 AA: 6644 MW: 19464 CN:
SQ
      Query Match 100.0%; Score 64; DB 10; Length 60;
      Best Local Similarity 100.0%; Pred. No. 2,93e-00;
      Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB      43 FIKRVSNVI 57
QY      1 FIKRVSNVI 9
      RESULT 15
      ID US-08-350-225-62 STANDARD; PRT: 60 AA.
      AC xxxxxx
      XX
      DT
      DE
      Sequence 62, Application US/08467697
      GENERAL INFORMATION:
      APPLICANT: Griffith, Irwin J.
      APPLICANT: Pollock, Joanne
      APPLICANT: Bond, Julian F.

```

Release 3.1A John F. Collins, Bioinformatics Research Unit,
Copyright (C) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run On: Mon Jun 19 16:06:27 2000; MasPar time 5.08 seconds
83.618 Million cell updates/sec

Tabular output not generated.

Title: >US-09-142-524A-7
Description: (1-9) from US09142524A.pep
Perfect Score: 64
Sequence: 1 FIKRVSNV: 9

Scoring table: BAW 150
Jsp 15

Searched: 142080 seqs, 47172406 residues

Post-Processing: Minimum Match 0%
Listing first 45 summaries

Database: pir62
1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 23.798; Variance 28.919; scale 0.823

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query Match | Length | DB ID | Description | Pred. No. |
|------------|-------------|--------|-------|---|-----------|
| 1 | 64 | 100.0 | 374 | major allergen Cry j 1 precursor (clone pCC1-15) - Japanese cedar | 8.03e-03 |
| 2 | 64 | 100.0 | 374 | major allergen Cry j 1 precursor (clone pCC1-15) - Japanese cedar | 8.03e-03 |
| 3 | 67 | 89.1 | 326 | petase lyase - Asper | 3.16e-01 |
| 4 | 51 | 79.7 | 720 | hypothetical protein | 5.98e-00 |
| 5 | 50 | 78.1 | 256 | beta-lactamase (EC 3.1.1.1) | 9.56e-00 |
| 6 | 50 | 78.1 | 317 | endo-glucanase - Therm | 9.56e-00 |
| 7 | 50 | 78.1 | 473 | hypothetical protein | 9.56e-00 |
| 8 | 50 | 78.1 | 479 | hypothetical protein | 9.56e-00 |
| 9 | 49 | 76.6 | 190 | transcription regulat | 1.52e-01 |
| 10 | 49 | 76.6 | 227 | beta-lactamase (EC 3.1.1.1) | 1.52e-01 |
| 11 | 49 | 76.6 | 227 | beta-lactamase (EC 3.1.1.1) | 1.52e-01 |
| 12 | 49 | 76.6 | 379 | probable membrane pro | 1.52e-01 |
| 13 | 49 | 76.6 | 379 | probable membrane pro | 1.52e-01 |
| 14 | 49 | 76.6 | 381 | probable membrane pro | 1.52e-01 |
| 15 | 49 | 76.6 | 432 | hypothetical protein | 1.52e-01 |
| 16 | 49 | 76.6 | 520 | cysteine--trna ligase | 1.52e-01 |
| 17 | 48 | 75.0 | 487 | carbamoyl-phosphate s | 2.39e-01 |
| 18 | 48 | 75.0 | 601 | probable glycosyl hyd | 2.39e-01 |
| 19 | 48 | 75.0 | 4725 | dynein heavy chain, c | 2.39e-01 |
| 20 | 47 | 73.4 | 234 | hypothetical protein | 3.75e-01 |
| 21 | 47 | 73.4 | 410 | 3-phosphoglycerate ki | 3.75e-01 |
| 22 | 47 | 73.4 | 511 | dihydrolipicamide S-su | 3.75e-01 |
| 23 | 47 | 73.4 | 537 | probable flagella-rel | 3.75e-01 |

24 47 73.4 899 2 C72285 hypothetical protein 3.75e-01
25 47 73.4 899 2 S49534 hypothetical protein 3.75e-01
26 47 73.4 1462 1 S32437 pol polyprotein - Vol 3.75e-01
27 46 71.9 143 2 A69115 hypothetical protein 5.92e-01
28 46 71.9 317 2 G70356 thymidylate synthase 5.92e-01
29 46 71.9 500 1 Q08E38 HHLF6 protein - human 5.82e-01
30 46 71.9 528 2 A37878 glucuronosyltransfera 5.82e-01
31 46 71.9 528 2 JN0619 JSP-glucuronosyltrans 5.82e-01
32 46 71.9 530 2 A38276 glucuronosyltransfera 5.82e-01
33 46 71.9 530 2 S07390 glucuronosyltransfera 5.82e-01
34 46 71.9 530 2 A42467 glucuronosyltransfera 5.82e-01
35 46 71.9 532 2 A25064 glucuronosyltransfera 5.82e-01
36 46 71.9 532 2 T16752 hypothetical protein 5.82e-01
37 46 71.9 545 1 JX0225 cytochrome P450 CYP19 5.82e-01
38 46 71.9 585 2 S73325 stage II sporulation 5.82e-01
39 46 71.9 585 2 S71018 stage II sporulation 5.82e-01
40 46 71.9 774 2 T14555 DNA polymerase holoic 5.82e-01
41 46 71.9 840 1 I43509 endopeptidase Clp AIP 5.82e-01
42 46 71.9 847 2 A48228 beta-N-acetylhexosami 5.64e-01
43 46 71.9 857 1 D35905 endopeptidase Clp AIP 5.82e-01
44 46 71.9 1006 2 C70445 Arpase subunit of ATP 5.82e-01
45 46 71.9 1274 2 T04018 hypothetical protein 5.82e-01

ALIGNMENTS

RESULT 1 JC2124 *type complete
ENTRY major allergen Cry j 1 precursor (clone pCC1-15) - Japanese cedar
TITLE
ORGANISM
DATE 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 26-Aug-1999

ACCESSIONS JC2124
REFERENCE JC2123
#authors Sore, T; Komiyama N; Shimizu, K; Kusakabe, T; Murakubo, K; Kino, K.
#journal Biochem. Biophys. Res. Commun. (1994) 199:519-525
#title Cloning and sequencing of cDNA coding for Cry j 1, a major allergen of Japanese cedar pollen.
#cross-references MIM:94163234
#accession JC2124
#molecule_type mRNA
#residues 1-374 #label S/N
#cross-references G0125459; N0149461; P01410687; P0141454
#experimental_data_Sources public
#note The authors describe a cDNA clone coding for Cry j 1, a major allergen of Japanese cedar pollen.

CLASSIFICATION #superfamily feature lyase (A1A)
KEYWORDS glycoprotein; pollen
FEATURE
1-21
22-374
158,191,293,354 #length 374 #molecular_weight 40702 #checksum 3592
#domain signal sequence #status predicted #label S/N
#product major allergen cry j 1 (clone pCC1-15) #status predicted #label MAT
#binding site carbohydrate (Asn) (covalent) #status predicted

SUMMARY
Query Match 100.0%; Score 64; DB 2; Length 374;
Best Local Similarity 100.0%; Prod. No. 8.03e-03;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 130 FIKRVSNV: 138
1111111111
QY 1 FIKRVSNV: 9

RESULT 2 JC2123 *type complete
ENTRY major allergen Cry j 1 precursor (clone pCC1-2-2) - Japanese cedar
TITLE
ORGANISM
DATE 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change

```

26-Aug-1999
ACCESSIONS   JC2123: PC2065
REFERENCE    Sure, T.; Koriyama, N.; Shimizu, K.; Kusakabe, T.; Morikubo,
#authors     K.; Kiro, K.
#journal:    Biochem. Biophys. Res. Commun. (1994) 199:619-625
#title       Cloning and sequencing of cDNA coding for Cry 1A, a major
#allergen of Japanese cedar pollen.
#cross-references MIM:34103234
#accession    JC2123
#molecule_type cDNA
#residues     1-374 #label SON
#cross-references GB:D26544; NID:q493631; PID:d1006086; PID:q493632
#experimental_source pollen
#accession    PC2065
#molecule_type protein
#residues     22-53;58-81;219-232;236-258;259-357;365-372 #label SC2
#note         the authors described carbohydrate binding site for
#             residue 279
#             glycoprotein: pollen.
CLASSIFICATION #superfamily pectate lyase LAT59
KEYWORDS        glycoprotein; pollen.
FEATURE
1-2:
22-374
159-191;273,354
#domain signal sequence #status predicted #label SIGN
#product major allergen Cry 1A (clone PC21-2-2); #status
#predicted #label MAIN
#binding_site carbohydrate (Asn) (covalent) #status
#predicted
SUMMARY         #length 374 #molecular-weight 40645 #checksum 2920
Query Match    100.0% Score 64: DB 2: Length 374:
Best Local Similarity 100.0% Pred. No. 8.03e-03:
Matches 9: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
Db 110 YKRVSNVI 138
#residues      1-374
QY 1 FIKRVSNVI 9
RESULT 3
ENTRY      SS1509 #type complete
#title      pectate lyase - Aspergillus sp.
#formal_name Aspergillus sp.
#journal    15-Jun-1995 #sequence_revision 15-Apr-1995 #text_change
#keywords   CREA, a regulator of carbon catabolite repression.
#feature     1-29
#cross-references MIM:D5308556
#accession   SS1509
#status      preliminary
#molecule_type DNA
#residues     1-326 #label HOM
SUMMARY      #length 326 #molecular-weight 34565 #checksum 3698
Query Match    89.1% Score 57: DB 2: Length 326:
Best Local Similarity 77.8% Pred. No. 3.16e-01:
Matches 7: Conservative 2: Mismatches 0: Indels 0: Gaps 0:
Db 110 YKRVSNVI 118
#residues      1-326
QY 1 FIKRVSNVI 9
RESULT 4
ENTRY      SS0920 #type complete
#title      hypothetical protein YKR192w - yeast (Saccharomyces
#formal_name Saccharomyces cerevisiae)
#journal    5-Sep-1999
#cross-references MIM:495646.04
#accession   SS0920
#status      hypothetical
#molecule_type protein
#residues     1-326
QY 1 FIKRVSNVI 9
#journal

```

```

10-Feb-1995 #sequence_revision 12-May-1995 #text_change
21-Nov-1997
ACCESSIONS   S50920
REFERENCE    Pearson, D.; Rowan, S.
#authors     Submitted to the EMBL Data Library, January 1995
#submission  S50920
#accession   S50920
#molecule_type DNA
#residues     1-720 #label PEA
#cross-references EMBL:Z47915; NID:3642280; PID:q642284; MIM:588192
GENETICS
#map_position 13R
SUMMARY      #length 720 #molecular-weight 82131 #checksum 676
Query Match    79.7% Score 51: DB 2: Length 720:
Best Local Similarity 55.6% Pred. No. 5.98e-06:
Matches 5: Conservative 4: Mismatches 0: Indels 0: Gaps 0:
Db 351 YKRVSNVI 359
#residues      1-720
QY 1 FIKRVSNVI 9
RESULT 5
ENTRY      A32017 #type complete
#title      beta-lactamase (EC 3.5.2.6) II precursor - Bacillus cereus
#formal_name Bacillus cereus
#journal    07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change
#keywords   cereus 5/B/6 beta-lactamase II structural gene.
#feature     1-279
#cross-references MIM:D8227875
#accession   A32017
#molecule_type DNA
#residues     1-256 #label IM
#cross-references GB:M19510; MIM:q143127; PID:AAA22562; PID:q143129
CLASSIFICATION #superfamily beta-lactamase II
KEYWORDS        antibiotic resistance; hydrolase
#feature     1-29
#domain signal sequence #status predicted #label SIGN
#product beta-lactamase (amase II) #status predicted #label MA
#journal    11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change
#accession   A32017
SUMMARY      #length 256 #molecular-weight 28749 #checksum 172
Query Match    79.7% Score 51: DB 2: Length 256:
Best Local Similarity 55.6% Pred. No. 5.98e-06:
Matches 5: Conservative 4: Mismatches 1: Indels 0: Gaps 0:
Db 104 FIKRVTDVI 112
#residues      1-256
QY 1 FIKRVSNVI 9
RESULT 6
ENTRY      B72216 #type complete
#title      endoglucanase - Thermotoga maritima (strain MSB8)
#formal_name Thermotoga maritima
#journal    11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change
#accession   B72216
#status      preliminary
#molecule_type protein
#residues     1-720
QY 1 FIKRVSNVI 9
#journal

```



```

title      Evidence for lateral gene transfer between Archaea and
            Bacteria from genome sequence of Thermotoga maritima.
cross-references MUID:9287316
accession    B72216
status      preliminary
molecule_type DNA
residues     1-317 #label ARN
cross-references GB:AE001613; GB:AE000512; NID:q4982323;
            TIGR:TM1751
experimental_source strain MSB8

GENETICS
gene        TM1751
summary     #length 317 #molecular-weight 37383 #checksum 6220

Query Match      78.1%; Score 50; DB 2; Length 317;
Best Local Similarity 66.7%; Pred. No. 9.56e+00;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

DB 72 FVKRVDEVI 90
      1 FIKRVSNVI 9

RESULT 7
ENTRY 7
TITLE 7
      B70541 #type complete
      hypothetical protein Rv1576c - Mycobacterium tuberculosis
      (strain H37RV)
ORGANISM #format_name Mycobacterium tuberculosis
DATE 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change
      17-Jul-1998
ACCESSIONS B70541
REFERENCE
authors Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher,
            C.; Harris, D.; Gordon, S.V.; Eigmeier, K.; Gas, S.; Barry
            III, C.E.; Tekala, F.; Badcock, K.; Basham, D.; Brown, D.;
            Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.;
            Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
            Hornsby, T.; Jagels, K.; Krogh, A.; McLean, J.; Mule, S.;
            Murphy, L.; Oliver, S.; Osborne, J.; Quail, M.A.;
            Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.;
            Skellton, S.; Squares, S.; Squires, R.; Sulston, J.E.;
            Taylor, K.; Whitehead, S.; Barrett, B.G.
            Nature (1998) 393:537-544
            Deciphering the biology of Mycobacterium tuberculosis from
            the complete genome sequence.
cross-references MUID:9285987
accession    B70541
status      preliminary; nucleic acid sequence not shown;
            translation not shown
molecule_type DNA
residues     1-479 #label COL
cross-references GB:Z80225; GB:AL123456; NID:q3242265; PID:q4981354;
            PID:q35069;
experimental_source strain H37RV

GENETICS
gene        Rv2650c
summary     #length 479 #mo. alar-weight 51099 #checksum 4376

Query Match      78.1%; Score 50; DB 2; Length 479;
Best Local Similarity 66.7%; Pred. No. 9.56e+00;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

DB 154 FVKRVSNPV 162
      1 FIKRVSNVI 9

RESULT 9
ENTRY 9
TITLE 9
      B72330 #type Complete
      transcription regulator, terR family - Thermotoga maritima
      (strain MSB6)
ORGANISM #format_name Thermotoga maritima
DATE 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change
      20-Aug-1999
ACCESSIONS B72330
REFERENCE
authors Nelson, K.R.; Clayton, R.A.; Gill, S.R.; Binkley, M.E.; Dodson,
            R.J.; Bhatt, D.B.; Hickey, E.K.; Peterson, J.D.; Nelson,
            W.C.; Ketchum, K.A.; McDonald, L.; Utterback, J.R.; Malek,
            J.A.; Linber, K.D.; Garrett, M.M.; Stewart, A.M.; Cotton,
            M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
            Heidelberg, J.; Sutton, G.G.; Fleischmann, R.D.; White, O.;
            Salzberg, S.L.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
            Nature (1999) 399:323-329
            Evidence for lateral gene transfer between Archaea and
            Bacteria from genome sequence of Thermotoga maritima.
cross-references MUID:9287316
accession    B72330
status      preliminary
molecule_type DNA
residues     1-190 #label ARN
cross-references GB:AE001749; GB:AE000512; NID:q4981346; PID:q4981354;
            TIGR:TM0823
experimental_source strain MSB8

GENETICS
gene        TM0823
summary     #length 190 #molecular-weight 22492 #checksum 7501

Query Match      76.6%; Score 49; DB 2; Length 190;
Best Local Similarity 66.7%; Pred. No. 1.52e+01;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

DB 148 FVKRVSNPV 156
      1 FIKRVSNVI 9

RESULT 8
ENTRY 8
TITLE 8
      F70965 #type complete
      hypothetical protein Rv2650c - Mycobacterium tuberculosis
      (strain H37RV)
ORGANISM #format_name Mycobacterium tuberculosis
DATE 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change

```

```

17-Jul-1998
F70965
A70500
Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher,
            C.; Harris, D.; Gordon, S.V.; Eigmeier, K.; Gas, S.; Barry
            III, C.E.; Tekala, F.; Badcock, K.; Basham, D.; Brown, D.;
            Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.;
            Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
            Hornsby, T.; Jagels, K.; Krogh, A.; McLean, J.; Mule, S.;
            Murphy, L.; Oliver, S.; Osborne, J.; Quail, M.A.;
            Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.;
            Skellton, S.; Squares, S.; Squires, R.; Sulston, J.E.;
            Taylor, K.; Whitehead, S.; Barrett, B.G.
            Nature (1998) 393:537-544
            Deciphering the biology of Mycobacterium tuberculosis from
            the complete genome sequence.
cross-references MUID:9285987
accession    F70965
status      preliminary; nucleic acid sequence not shown;
            translation not shown
molecule_type DNA
residues     1-479 #label COL
cross-references GB:Z80225; GB:AL123456; NID:q3242265; PID:q4981354;
            PID:q35069;
experimental_source strain H37RV

GENETICS
gene        Rv2650c
summary     #length 479 #mo. alar-weight 51099 #checksum 4376

Query Match      78.1%; Score 50; DB 2; Length 479;
Best Local Similarity 66.7%; Pred. No. 9.56e+00;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

DB 154 FVKRVSNPV 162
      1 FIKRVSNVI 9

RESULT 9
ENTRY 9
TITLE 9
      B72330 #type Complete
      transcription regulator, terR family - Thermotoga maritima
      (strain MSB6)
ORGANISM #format_name Thermotoga maritima
DATE 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change
      20-Aug-1999
ACCESSIONS B72330
REFERENCE
authors Nelson, K.R.; Clayton, R.A.; Gill, S.R.; Binkley, M.E.; Dodson,
            R.J.; Bhatt, D.B.; Hickey, E.K.; Peterson, J.D.; Nelson,
            W.C.; Ketchum, K.A.; McDonald, L.; Utterback, J.R.; Malek,
            J.A.; Linber, K.D.; Garrett, M.M.; Stewart, A.M.; Cotton,
            M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
            Heidelberg, J.; Sutton, G.G.; Fleischmann, R.D.; White, O.;
            Salzberg, S.L.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
            Nature (1999) 399:323-329
            Evidence for lateral gene transfer between Archaea and
            Bacteria from genome sequence of Thermotoga maritima.
cross-references MUID:9287316
accession    B72330
status      preliminary
molecule_type DNA
residues     1-190 #label ARN
cross-references GB:AE001749; GB:AE000512; NID:q4981346; PID:q4981354;
            TIGR:TM0823
experimental_source strain MSB8

GENETICS
gene        TM0823
summary     #length 190 #molecular-weight 22492 #checksum 7501

Query Match      76.6%; Score 49; DB 2; Length 190;
Best Local Similarity 66.7%; Pred. No. 1.52e+01;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```



```

Best Local Similarity 55.6%  Pred. No. 1.52e+01:
Matches 5: Conservative 4: Mismatches 0: Indels 0: Gaps 0:

DB 229 FLKRISNIF 237
    |||:||||
    1 FIKRVSNVI 9

RESULT 13
ENTRY
TITLE
ALTERNATE_NAMES
ORGANISM
DATE
ACCESSIONS
REFERENCE
AUTHORS
SUBMISSION
ACCESSION
MOLECULE_TYPE DNA
RESIDUES 1-47 #label PEL
CROSS-REFERENCES EMBL:236170; MIPS:YBR302c
EXPERIMENTAL_SOURCE strain S288C
GENETICS C02
REFERENCE
AUTHORS
SUBMISSION
ACCESSION
MOLECULE_TYPE DNA
RESIDUES 30-379 #label FEU
CROSS-REFERENCES EMBL:236170; MIPS:YBR302c
EXPERIMENTAL_SOURCE strain S288C
GENETICS C02
REFERENCE
AUTHORS
SUBMISSION
ACCESSION
MOLECULE_TYPE DNA
RESIDUES 30-379 #label FEU
CROSS-REFERENCES EMBL:236170; MIPS:YBR302c
EXPERIMENTAL_SOURCE strain S288C
GENETICS C02
REFERENCE
AUTHORS
SUBMISSION
ACCESSION
MOLECULE_TYPE DNA
RESIDUES 1-379 #label BOW
CROSS-REFERENCES EMBL:250178; NID:G927528; PIDN:CAA90551.1;
PID:G927530; MIPS:YML132w
EXPERIMENTAL_SOURCE strain AB972
GENETICS C02
AUTHORS
SUBMISSION
ACCESSION
MOLECULE_TYPE DNA
RESIDUES 1-379 #label BOW
CROSS-REFERENCES MIPS:YML132w; SGD:S0004601
MAP_POSITION 13L
CLASSIFICATION
KEYWORDS
FEATURE
domain transmembrane #status predicted #label TM1\
domain transmembrane #status predicted #label TM2\
domain transmembrane #status predicted #label TM3\
233-253

```

```

254-273
SUMMARY
domain transmembrane #status predicted #label TM4
length 379 #molecular-weight 45165 #checksum 1211

Query Match 76.6%  Score 49: DB 2: Length 379:
Best Local Similarity 55.6%  Pred. No. 1.52e+01:
Matches 5: Conservative 4: Mismatches 0: Indels 0: Gaps 0:

DB 229 FLKRISNIF 237
    |||:||||
    1 FIKRVSNVI 9

RESULT 14
ENTRY
TITLE
ALTERNATE_NAMES
ORGANISM
DATE
ACCESSIONS
REFERENCE
AUTHORS
SUBMISSION
ACCESSION
MOLECULE_TYPE DNA
RESIDUES 1-381 #label OBE
CROSS-REFERENCES EMBL:271611; NID:q1302452; PIDN:CAA94270.1;
PID:q23973; PID:q1302453; MIPS:YNL336w
EXPERIMENTAL_SOURCE strain S288C
GENETICS
AUTHORS
SUBMISSION
ACCESSION
MOLECULE_TYPE DNA
RESIDUES 1-381 #label OBE
CROSS-REFERENCES SGD:S0005; MIPS:YNL336w
MAP_POSITION 14L
CLASSIFICATION
KEYWORDS
FEATURE
domain transmembrane #status predicted #label TM1\
domain transmembrane #status predicted #label TM2\
domain transmembrane #status predicted #label TM3\
length 381 #molecular-weight 45288 #checksum 6842

Query Match 76.6%  Score 49: DB 2: Length 381:
Best Local Similarity 55.6%  Pred. No. 1.52e+01:
Matches 5: Conservative 4: Mismatches 0: Indels 0: Gaps 0:

DB 229 FLKRISNIF 237
    |||:||||
    1 FIKRVSNVI 9

RESULT 15
ENTRY
TITLE
ORGANISM
DATE
ACCESSIONS
REFERENCE
AUTHORS
GENETICS
AUTHORS
SUBMISSION
ACCESSION
MOLECULE_TYPE DNA
RESIDUES 1-379 #label BOW
CROSS-REFERENCES EMBL:250178; NID:G927528; PIDN:CAA90551.1;
PID:G927530; MIPS:YML132w
EXPERIMENTAL_SOURCE strain AB972
GENETICS C02
AUTHORS
SUBMISSION
ACCESSION
MOLECULE_TYPE DNA
RESIDUES 1-379 #label BOW
CROSS-REFERENCES MIPS:YML132w; SGD:S0004601
MAP_POSITION 13L
CLASSIFICATION
KEYWORDS
FEATURE
domain transmembrane #status predicted #label TM1\
domain transmembrane #status predicted #label TM2\
domain transmembrane #status predicted #label TM3\
233-253

```

```
*accession G64403
**status preliminary: nucleic acid sequence not shown:
**translation not shown
**molecule_type DNA
**residues 1-432 **label BUL
**cross-references GB:U57526; GB:L27117; NID:9151508; PDB:gl591518;
               TIGR:MJ0831; PDB:gl51089
GENETICS
*map_position FOR750952-752250
SUMMARY *length 432 *molecular_weight 48835 *checksum 5047
Query Match 76.68; Score 49; DB 2; Length 432;
Best Local Similarity 55.68; Pred. No. 1; Size 0;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
DB 418 FCKRVDE:: 426
QY 1 FCKRVSNVI 9
```

Search completed: Mon Jun 19 16:06:33 2000
Job time : 6 secs.

 W E B

 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Mrsrch_pp protein - protein database search, using Smith-Waterman algorithm
 Run On: Mon Jun 19 16:05:39 2000; MasPar time 3.40 Seconds
 Tucleur output not generated; 80,554 Million cell updates/sec

Title: US-09-142-524A-7
 Description: (1-9) from US9142524A.ref
 Perfect Score: 64
 Sequence: : FIKRVSNVI 9
 Scoring table: PAM 150
 Gap 15

Searched: 81857 seqs, 30454973 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: swiss-prot38
 l:swissprot

Statistics: Mean 24.343; Variance 25.165; scale 1.967

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Length | DB ID | Description | Pred. No. |
|------------|-------|--------------|-------|--------------------------|-----------|
| 1 | 64 | 100 | 0 | SUGI BASIC PROTEIN PRE | 6.81e-04 |
| 2 | 57 | 89.1 | 326 | PECTATE LYASE PRECURSO | 4.57e-02 |
| 3 | 51 | 79.7 | 720 | HYPOTHETICAL 82.1 KD P | 1.36e-00 |
| 4 | 50 | 78.1 | 129 | CYTOSTATIN PRECURSOR (OV | 2.32e-00 |
| 5 | 50 | 78.1 | 255 | BETA-LACTAMASE, TYPE I | 2.32e-00 |
| 6 | 49 | 75.6 | 257 | BETA-LACTAMASE, TYPE I | 3.94e-00 |
| 7 | 49 | 75.6 | 257 | BETA-LACTAMASE, TYPE I | 3.94e-00 |
| 8 | 49 | 75.6 | 379 | HYPOTHETICAL 45.3 KD P | 3.94e-00 |
| 9 | 49 | 75.6 | 379 | HYPOTHETICAL 45.3 KD P | 3.94e-00 |
| 10 | 49 | 75.6 | 381 | HYPOTHETICAL 45.3 KD P | 3.94e-00 |
| 11 | 49 | 75.6 | 432 | HYPOTHETICAL 103.0 KD | 3.94e-00 |
| 12 | 49 | 75.6 | 520 | CYSTEINYL-TRNA SYNTHET | 3.94e-00 |
| 13 | 49 | 75.6 | 501 | HYPOTHETICAL 68.8 KD P | 3.94e-00 |
| 14 | 48 | 75.0 | 288 | 4-HYDROXYBENZATE OCTA | 6.62e-00 |
| 15 | 48 | 75.0 | 4725 | DYNEIN HEAVY CHAIN, CY | 6.62e-00 |
| 16 | 47 | 73.4 | 509 | PYRUVATE KINASE (EC 2. | 1.10e-01 |
| 17 | 47 | 73.4 | 789 | LARVAL SERUM PROTEIN 1 | 1.10e-01 |
| 18 | 47 | 73.4 | 899 | HYPOTHETICAL 103.0 KD | 1.10e-01 |
| 19 | 46 | 71.9 | 425 | ADENOSYLMETHIONINE-8-A | 1.82e-01 |
| 20 | 46 | 71.9 | 500 | HYPOTHETICAL PROTEIN H | 1.82e-01 |
| 21 | 46 | 71.9 | 528 | UDP-GLUCURONOSYLTRANSF | 1.82e-01 |
| 22 | 46 | 71.9 | 530 | UDP-GLUCURONOSYLTRANSF | 1.82e-01 |
| 23 | 46 | 71.9 | 530 | UDP-GLUCURONOSYLTRANSF | 1.82e-01 |

| RESULT ID | SBP_CRYJA | STANDARD | PRT | 374 AA |
|-----------|--|----------|-----|--------|
| AC | P18632 | | | |
| DI | 01-NOV-1990 (Rel. 16, Created) | | | |
| DI | 01-NOV-1995 (Rel. 32, Last sequence update) | | | |
| DI | 15-DEC-1998 (Rel. 37, Last annotation update) | | | |
| DE | SUGI: BASIC PROTEIN PRECURSOR (SBP) (MAJOR ALLERGEN CRY J 1) (CRY J 1) | | | |
| OS | Cryptomeria japonica (Japanese cedar) | | | |
| CC | Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphylloytes; Spermatophyta; Coniferopsida; Coniferales; Taxodiaceae; Cryptomeria | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A., AND PARTIAL SEQUENCE | | | |
| RC | TISSUE=POLLEN | | | |
| RX | MEDLINE: 94183234 | | | |
| RA | Sone T., Komiyama N., Shimizu K., Kusakabe T., Morikubo K., Kino K., | | | |
| RA | Cloning and sequencing of cDNA coding for Cry J 1, a major allergen of Japanese cedar pollen. | | | |
| RI | of Japanese cedar pollen. | | | |
| RI | Biochem. Biophys. Res. Commun. 199 199:251-254 | | | |
| RN | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | TISSUE=POLLEN | | | |
| RA | Narba M., Karose M., Torigoe K., Fukuda S., Kurimoto M., Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases. | | | |
| RN | [3] | | | |
| RP | SEQUENCE OF 22-41 | | | |
| RC | TISSUE=POLLEN | | | |
| RX | MEDLINE: 89031257 | | | |
| RA | Tanai M., Ando S., Usui M., Kurimoto M., Sakaguchi M., Inouye S., Natuhasi T., | | | |
| RA | "N-terminal amino acid sequence of a major allergen of Japanese cedar pollen (Cry j 1)." | | | |
| RI | FEBS Lett. 239:329-332(1988). | | | |
| RN | [4] | | | |
| RP | CARBOHYDRATES | | | |
| RX | MEDLINE: 95003748 | | | |
| RA | Hijikata A., Matsumoto I., Kojima K., Ogawa H., | | | |
| RA | "Antigenicity of the oligosaccharide moiety of the Japanese cedar (Cryptomeria japonica) pollen allergen, Cry j 1." | | | |
| RI | Int. Arch. Allergy Immunol. 105:198-202(1994) | | | |
| CC | 1- PTM: CONTAINS FUCOSE/XULOSE-CONTAINING N-LINKED OLIGOSACCHARIDES. | | | |
| CC | 1- DISEASE: THIS PROTEIN IS THE MAJOR ALLERGEN OF JAPANESE CEDAR POLLEN, THE MOST COMMON POLLEN ALLERGEN IN JAPAN. | | | |
| CC | 1- MISCELLANEOUS: THE SEQUENCE OF CRY J 1 FORM A IS SHOWN HERE. FORM B DIFFERS IN SIX POSITIONS. | | | |
| CC | 1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1. | | | |

AMB A I/AMB A II/CRY J I SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DE EMBL: D26544; BAA0542.1; 1
 DE EMBL: D26545; BAA0543.1; 1
 DE EMBL: D34639; BAA0702.1; 1
 DE EMBL: D44773; A44773
 DE EMBL: PF00544; pec-lyase; 1
 DE PRIN: PR00801; AMBALLEGEN
 KW Allergen; Glycoprotein; Multigene family; Signal;
 FT SIGNAL 1 21
 FT CHAIN 22 374 SUGI BASIC PROTEIN.
 FT VARIANT 12 12 L -> F (IN CRY J 1-B).
 FT VARIANT 143 143 H -> Y (IN CRY J 1-B).
 FT VARIANT 202 202 S -> T (IN CRY J 1-B).
 FT VARIANT 221 221 L -> S (IN CRY J 1-B).
 FT VARIANT 358 358 Q -> H (IN CRY J 1-B).
 FT VARIANT 361 361 K -> Q (IN CRY J 1-B).
 FT VARIANT 158 158 POTENTIAL.
 FT CARBOHYD 191 191 POTENTIAL.
 FT CARBOHYD 293 293 POTENTIAL.
 SC SEQUENCE 374 AA: 40645 MW: 74AB25950248F56F CR644;
 Query Match 100.0% Score 64 DB 1 Length 374;
 Best Local Similarity 100.0% Pred. No. 5.8.e-04;
 Matches 9: Conservative 0: Mismatches 0: Indels 0: Gaps 0;

DE 130 FIKRVSNVI 128
 QY 1 FIKRVSNVI 9

RESULT 2
 ID PEL ENVI STANDARD: PRT: 325 AA.
 AC Q00645;
 DI 15-DEC-1998 (Rel. 37, Created);
 DI 15-DEC-1998 (Rel. 37, Last sequence update);
 DI 15-DEC-1998 (Rel. 37, Last annotation update);
 DE PECLATE LYASE PRECURSOR (P12.2.2.2);
 GN PELA.
 OS Emricella nidulans (Aspergillus nidulans).
 OS Eukaryota; Fungi; Ascomycota; Plectomycetes; Emericellales;
 OS Trichocomaceae; Emericella.
 RN [1]
 RE SEQUENCE FROM N.A.
 RC STRAIN=F05C 4;
 RA MEDLINE: 95308536.
 RA Ho M.C., Whitehead M.P., Cleveland T.E., Dean R.A.;
 R1 "Sequence analysis of the Aspergillus nidulans peclate lyase pela
 R1 gene and evidence for binding of promoter regions to CRE, a
 R1 regulator of carbon catabolite repression.";
 R1 Curr. Genet. 27:142-149(1995).
 CC -!- CATALYTIC ACTIVITY: ELIMINATIVE CLEAVAGE OF PECTATE TO GIVE
 CC OLIGOSACCHARIDES WITH 4-DEOXY-ALPHA-D-MANN-4-ENURONOSYL GROUPS
 CC AT THEIR NON-REDUCING ENDS.
 CC -!- SURCELLULAR LOCATION: SECRETED (POTENTIAL).
 CC -!- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: U05592; AAA80568.1; 1
 DR PRAM: PF00544; pec-lyase; 1
 KW Lyase; Signal;
 FT SIGNAL 1 15 POTENTIAL.
 FT CHAIN 16 326 PECTATE LYASE.
 SQ SEQUENCE 326 AA: 34580 MW: F6445A4A63615049 CR064;
 Query Match 89.1% Score 57 DB 1 Length 126;
 Best Local Similarity 77.9% Pred. No. 4.67e-02;
 Matches 7: Conservative 2: Mismatches 0: Indels 0: Gaps 0;

DE 110 YIKRVSNVI 118
 QY 1 FIKRVSNVI 9

RESULT 3
 ID YMS2-YEAST STANDARD: PRT: 720 AA.
 AC Q04322;
 DI 01-NOV-1997 (Rel. 35, Created);
 DI 01-NOV-1997 (Rel. 35, Last sequence update);
 DI 01-NOV-1997 (Rel. 35, Last annotation update);
 DE HYPOTHETICAL 82.1 KD PROTEIN IN S35-MRP24 INTERGENIC REGION
 GN YMR12W OR YMR646.04.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OS Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
 OS Saccharomycetaceae; Saccharomyces.
 RN [1]
 RE SEQUENCE FROM N.A.
 RC STRAIN=S288C/AB972;
 RA Pearson C., Bowman S., Barrell B.G., Bolandream M.A.;
 R1 Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: TO YEAST YPL249C AND S.POMBE SPAC28F1.09.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: 247815; CAA87813.1; 1
 KW Hypothetical protein;
 SQ SEQUENCE 720 AA: 82131 MW: 8142064F587247 CR764;
 Query Match 79.7% Score 51 DB 1 Length 720;
 Best Local Similarity 55.4% Pred. No. 1.66e-01;
 Matches 5: Conservative 4: Mismatches 0: Indels 0: Gaps 0;

DE 351 YIKRVSNVI 359
 QY 1 FIKRVSNVI 9

RESULT 4
 ID CYI-CYCPA STANDARD: PRT: 125 AA.
 AC P35481;
 DI 01-JUN-1994 (Rel. 29, Created);
 DI 01-JUN-1994 (Rel. 29, Last sequence update);
 DI 15-DEC-1998 (Rel. 37, Last annotation update);
 DE CYSTATIN PRECURSOR (OVARIAN CYSTATIN) (P12).
 OS Cyprinus carpio (Common carp).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Actinopterygii;
 OS Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes;
 OS Cyprinidae; Cyprininae; Cyprinus.
 RN [1]
 RE SEQUENCE FROM N.A., AND SEQUENCE OF 19-57, 74-80 AND 92-129.
 CC -!- ISSUE-OVARY;
 RX MEDLINE: 96208938.
 RA Tsai Y.-J., Chang G.-D., Huang C.-J., Chang Y.-S., Huang F.-L.;
 R1 "Purification and molecular cloning of carp ovarian cystatin.";
 R1 Comp. Biochem. Physiol. 113B:573-580(1996).

CC -1- FUNCTION: CYSTEINE PROTEINASE INHIBITOR.
 CC -1- TISSUE SPECIFICITY: DETECTED IN OVARY BUT NOT IN ALL THE OTHER
 CC EXAMINED TISSUES.
 CC -1- PIM: THIS PROTEIN IS FURTHER PROCESSED BY AN INTRAMOLECULAR
 CC CLEAVAGE TO PRODUCE TWO CHAINS LINKED BY A DISULFIDE BRIDGE.
 CC -1- SIMILARITY: BELONGS TO THE TYPE 2 CYSTATIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: L23572; AAB42011.1;
 CC HSSP: P00398; I490.
 CC PROSITE: PS00287; CYSTATIN_1.
 CC PRAM: PF00031; cystatin_1.
 CC Tricol protease inhibitor: Signal.
 CC SIGNAL 1 18
 CC CHAIN 19 129
 CC ACT_SITE 23 23
 CC REACTIVE SITE (BY SIMILARITY).
 CC SECONDARY AREA OF CONTACT.
 CC SITE 57 71
 CC BY SIMILARITY.
 CC DISULFID 85 94
 CC BY SIMILARITY.
 CC SITE 108 128
 CC CLEAVAGE (PROBABLE).
 CC SITE 91 92
 CC SEQUENCE 129 AA: 14236 MW: 732760DF347515BC CRC64;
 CC
 CC Query Match 78.1% Score 50; DB 1; Length 129;
 CC Best Local Similarity 55.6%; Pred. No. 232e-00;
 CC Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC 54 FVKRVSKVI 62
 CC 1:1:1:1:
 CC 1 FIKRVSNVI 9
 CC
 CC RESULT 5
 CC ID BL22BACCE STANDARD: PRT: 256 AA.
 CC AC P14498;
 CC DT 01-JAN-1990 (Rel. 13, Created)
 CC DI 03-JAN-1990 (Rel. 13, last sequence update)
 CC D1 15-JUL-1999 (Rel. 38, last annotation update)
 CC DE BETA-LACTAMASE, TYPE 1; PROTEINER (B: 3.5.2.6) (PENICILLINASE)
 CC (CEPHALOSPORINASE).
 CC OS Bacillus cereus
 CC OC Bacteria; Firmicutes; Bacillus/Clostridium group.
 CC ON Bacillus/Staphylococcus group; Bacillus.
 CC [1]
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE: 88227875,
 CC PA Lim H.M., Pene J.J., Shaw R.W.,
 CC R1 "Cloning, nucleotide sequence, and expression of the Bacillus cereus
 CC 5/5/6 beta-lactamase II structural gene."
 CC RL J. Bacteriol. 170:2873-2878(1988).
 CC [2]
 CC RP MUTAGENESIS
 CC RX MEDLINE: 93080559,
 CC RA Lim H.M., Pene J.J.,
 CC R1 "Mutations affecting the catalytic activity of Bacillus cereus 5/5/6
 CC beta-lactamase II."
 CC PL J. Biol. Chem. 264:11682-11687(1989).
 CC [3]
 CC RP MUTAGENESIS
 CC RX MEDLINE: 91264793,
 CC RA Lim H.M., Iyer R.K., Pene J.J.,
 CC R1 "Site-directed mutagenesis of dicarboxylic acids near the active site
 CC of Bacillus cereus 5/5/6 beta-lactamase II."
 CC RL Biochem. J. 276:403-404(1991).
 CC -1- FUNCTION: CAN HYDROLYZE CARAPENEM COMPOUNDS.
 CC -1- CATALYTIC ACTIVITY: A BETA-LACTAM + H(2)O -> A SUBSTITUTED BETA-

CC AMINO ACID.
 CC -1- COFACTOR: BINDS TWO ZINC IONS PER MOLECULE.
 CC -1- SIMILARITY: BELONGS TO THE CLASS-B BETA-LACTAMASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M19530; AAA22562.1;
 CC PIR: A32017; A32017.
 CC HSSP: P04190; IBM.
 CC PROSITE: PS00743; BETA-LACTAMASE_B_1;
 CC PRAM: PF00753; Lactamase_B_1.
 CC Hydrolase; zinc; Antibiotic resistance. Signal.
 CC SIGNAL 1 29
 CC CHAIN 30 256
 CC METAL 115 115
 CC ZINC 1 (BY SIMILARITY).
 CC METAL 117 117
 CC ZINC 1 (BY SIMILARITY).
 CC METAL 119 119
 CC ZINC 2 (BY SIMILARITY).
 CC METAL 178 178
 CC ZINC 1 (BY SIMILARITY).
 CC METAL 197 197
 CC ZINC 2 (BY SIMILARITY).
 CC METAL 239 239
 CC ZINC 2 (BY SIMILARITY).
 CC MUTAGEN 110 110
 CC D-N: NO CHANGE IN ACTIVITY.
 CC MUTAGEN 119 119
 CC D-N: INACTIVATES THE ENZYME.
 CC MUTAGEN 177 177
 CC G-N: INACTIVATES THE ENZYME.
 CC SEQUENCE 256 AA: 28038 MW: 670F52378C35C2D CRC64;
 CC
 CC Query Match 78.1% Score 50; DB 1; Length 256;
 CC Best Local Similarity 66.7%; Pred. No. 2.32e-00;
 CC Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC 104 FKKRVTDVI 112
 CC 1:1:1:1:
 CC 1 FIKRVSNVI 9
 CC
 CC RESULT 6
 CC ID BL22BACCE STANDARD: PRT: 257 AA.
 CC AC P04190;
 CC DT 20-MAR-1987 (Rel. 14, Created)
 CC DI 20-MAR-1987 (Rel. 14, last sequence update)
 CC D1 15-FEB-2000 (Rel. 39, last annotation update)
 CC DE BETA-LACTAMASE, TYPE 1; PROTEINER (B: 3.5.2.6) (PENICILLINASE)
 CC (CEPHALOSPORINASE).
 CC ON Bacillus cereus
 CC OC Bacteria; Firmicutes; Bacillus/Clostridium group.
 CC ON Bacillus/Staphylococcus group; Bacillus.
 CC [1]
 CC RP SEQUENCE FROM N.A.
 CC RX STRAIN: 569/H.
 CC RX MEDLINE: 86008056,
 CC RA Hussain M., Carlino A., Madona M.J., Lampen J.O.,
 CC R1 "Cloning and sequencing of the metalloprotein beta-lactamase II
 CC gene of Bacillus cereus 569/H in Escherichia coli."
 CC RL J. Bacteriol. 164:223-229(1985).
 CC [2]
 CC RP SEQUENCE OF 31-183; 187-2. AND 214-257.
 CC RX STRAIN: 569/H.
 CC RX MEDLINE: 86005446,
 CC RA Ambler R.P., Daniel M., Fleming J., Hermoso J.M., Pang C., Watley S.G.,
 CC R1 "The amino acid sequence of the zinc-requiring beta-lactamase II from
 CC the bacterium Bacillus cereus 569/H."
 CC RL FEBS Lett. 189:207-211(1985).
 CC [3]
 CC RP X-RAY CRYSTALLOGRAPHY (3.5 ANGSTROMS).
 CC RX MEDLINE: 9813384,
 CC RA Sutton B.J., Artymus P.J., Cordeiro-Borboa A.E., Little G.,

```

Matches      6;   Conservative      Mismatches      1;   Indels      0;   Gaps      0;

Db      105 FOKRVIDW1 113
QY      1 FIKRVSNV1 9

RESULT      7
ID      BLAZ_BACSP      STANDARD:      PRT:      457 AA.
DI      P10425:
DT      01-MAR-1995 (Rel. 10, Created)
DT      01-MAR-1995 (Rel. 10, Last sequence update)
DI      15-JUL-1999 (Rel. 38, Last annotation update)
DE      BETA-LACTAMASE, TYPE II PROTEIN A (EC 3.5.2.6) (PENICILLINASE)
OS      Bacillus sp. (strain 170)
OC      Bacteria; Firmicutes; Bacilli; Clostridiaceae group;
OC      Bacillus/staphylococcus group; Bacillus
RN      ..
RP      SEQUENCE FROM N.A., AND SEQUENCE OF XI-48
RX      MEDLINE: 86170399.
RT      Kato C., Kudo T., Watanabe K., Horikoshi K.
RI      Nucleotide sequence of the beta-lactamase gene of alkalophilic
RI      Bacillus sp. strain 170.
RL      J. Gen. Microbiol. 131:331-3324 (1985).
CC      FUNCTION: CAN HYDROLYZE CARBAPENEM COMPOUNDS.
CC      -| CATALYTIC ACTIVITY: A BETA-LACTAM - H(2)O + A SUBSTITUTED BETA-
CC      AMINO ACID.
CC      -| COFACTOR: BINDS TWO ZINC IONS PER MOLECULE.
CC      -| SIMILARITY: BELONGS TO THE CLASS-B BETA-LACTAMASE FAMILY
CC      -----
CC      This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL database at
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isc-sib.ch/announcements)
CC      or send an email to license@sib-switzerland.ch.
CC      -----
DB      EMBL: M15350; AAA22275.1;
DR      PIR: A24393; PNB525.
DR      HSSP: P04190; LBMC.
DR      PROSITE: PS00743; BETA_LACTAMASE_B_1;
DR      PROSITE: PS00744; BETA_LACTAMASE_B_2;
DR      PFAM: PF00753; lactamase_B;
KW      Hydrolase; Zinc; Antibiotic; Inhibitor; Signal;
FT      SIGNAL          1..30
FT      CHAIN           1..257
FT      METAL           116 1 (BY SIMILARITY)
FT      METAL           118 1 (BY SIMILARITY)
FT      METAL           120 1 (BY SIMILARITY)
FT      METAL           179 1 (BY SIMILARITY)
FT      METAL           198 1 (BY SIMILARITY)
FT      METAL           240 2 (BY SIMILARITY)
SQ      SEQUENCE      257 AA: 2815 MW: 234680 Da: FEEDDCG45AE571 CRG64;

Query Match      76.6%; Score 49; DB 1; Length 257;
Best Local Similarity 66.7%; Pred.No. 3,94e+06;
Matches      6;   Conservative      Mismatches      2;   Indels      0;   Gaps      0;

Db      105 FOKRVIDW1 113
QY      1 FIKRVSNV1 9

RESULT      8
ID      YFG2_YEAST      STANDARD:      PRT:      379 AA.
DI      P43542:
DT      01-NOV-1995 (Rel. 32, Created)
DT      01-NOV-1995 (Rel. 32, Last sequence update)
DI      05-NOV-1997 (Rel. 35, Last annotation update)
DE      HYPOTHETICAL 45.3 KD PROTEIN IN THIS 5'REGION.
OS      Saccharomyces cerevisiae (Baker's yeast).
OC      GN      YFL062W.

```


*Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.

[1]
SEQUENCE FROM N.A.
MEDLINE: 95400292.
RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,
RA Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
RA Yamazaki M., Tashiro H., Eki T.,
R *Analysis of the nucleotide sequence of chromosome VI from
R Saccharomyces cerevisiae.
R Nat. Genet. 10:251-258(1995).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE DUP FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

EMBL: D50612; BAA09179.1; -
EMBL: D45633; BAA08057.1; -
DR PFAM: PF00674; DUP; 2.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 43 63 POTENTIAL.
FT TRANSMEM 70 90 POTENTIAL.
FT TRANSMEM 233 253 POTENTIAL.
FT TRANSMEM 255 275 POTENTIAL.
SQ SEQUENCE 379 AA: 45312 MW: 522356.969890275 CRC64:

Query Match 76.6% Score 49: DB 1: Length 379;
Best Local Similarity 55.6% Pred. No. 3,94e+00;
Matches 5: Conservative 4: Mismatches 0: Indels 0: Gaps 0;

DB 229 FLKRISNIF 237

QY 1 FIKRVSNI 9

RESULT 9
ID YB76-YEAST STANDARD PRI: 479 AA.
AC P38363.
DE 01-OCT-1994 (Rel. 30, Created)
DE 01-OCT-1994 (Rel. 30, Last sequence update)
DE 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHETICAL 45.2 KD PROTEIN IN WAL35 3'REGION AND IN COX14 5'REGION.
GN (YBR302C OR YBR2121) AND (YML132W OR YML987.03).
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
OC Saccharomycetaceae; Saccharomyces.
RN [1]
RP SEQUENCE OF 1-47 FROM N.A. (YBR302C).
RC STRAIN-S288C.
RA Feldmann H., Mannhaupt G., Schwarze C., Vetter J.;
RC Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 30-379 FROM N.A. (YBR302C).
RC STRAIN-S288C.
RA Feldmann M., Potier S., Souciet J.-L.;
RN Submitted (AUG-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A. (YML132W).
RA Bowman S., Louis E.J., Barrell B.G., Rajandream M.A., Walsh S.V.;
RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE DUP FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

EMBL: Z36271; CAA85267.1; -
EMBL: Z50178; CAA90551.1; -
DR PIR: S46187; S46187.
DR PFAM: PF00674; DUP; 2.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 73 93 POTENTIAL.
FT TRANSMEM 255 275 POTENTIAL.
SQ SEQUENCE 379 AA: 45165 MW: 2AAB51B6C104AF61 CRC64:

Query Match 76.6% Score 49: DB 1: Length 379;
Best Local Similarity 55.6% Pred. No. 3,94e+00;
Matches 5: Conservative 4: Mismatches 0: Indels 0: Gaps 0;

DB 229 FLKRISNIF 237

QY 1 FIKRVSNI 9

RESULT 10
ID YN76-YEAST STANDARD PRI: 391 AA.
AC P53822.
DE 01-OCT-1996 (Rel. 34, Created)
DE 01-OCT-1996 (Rel. 34, Last sequence update)
DE 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHETICAL 45.3 KD PROTEIN IN THL12 5'REGION.
GN YNL336W OR N0275.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetaceae;
OC Saccharomycetaceae; Saccharomyces.
RN [1]
RP SEQUENCE FROM N.A.
RA Obermaier B., Piravandi E., Rinke M.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE DUP FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC

EMBL: Z71612; CAA96270.1; -
DR PFAM: PF00674; DUP; 2.
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 43 63 POTENTIAL.
FT TRANSMEM 73 93 POTENTIAL.
FT TRANSMEM 232 252 POTENTIAL.
FT TRANSMEM 255 275 POTENTIAL.
SQ SEQUENCE 381 AA: 45288 MW: F9AF930CB1935131 CRC64:

Query Match 76.6% Score 49: DB 1: Length 381;
Best Local Similarity 55.6% Pred. No. 3,94e+00;
Matches 5: Conservative 4: Mismatches 0: Indels 0: Gaps 0;

DB 229 FLKRISNIF 237

QY 1 FIKRVSNI 9

RESULT 11
ID Y831-METJA STANDARD PRI: 432 AA.
AC Q58241.
DE 01-NOV-1997 (Rel. 35, Created)
DE 01-NOV-1997 (Rel. 35, Last sequence update)

CC 11- SUBUNIT: MONOMER (BY SIMILARITY).

CC 11- SUBCELLULAR LOCATION: CYTOPLASMIC.

CC 11- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY

CC STRONG, TO METHIONYL-TRNA SYNTHETASE.

CC

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL institution

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to license@isb-sib.ch).

CC

CC EMBL: A5001193; AAC65055.1

CC TIGR: TP0091

CC PRINTS: PR05983; TRNASYNTHCYS

CC PROSITE: PS0078; AA-TRNA-LIGASE_I; FALSE_NEG.

CC PFAM: PF01426; tRNA-synthetase; 1

CC Aminoacyl-TRNA synthetase: Protein biosynthesis: class: ATP-binding

CC SIMILAR 3: 41

CC SIMILAR 30: 305 "HIGH" REGION

CC BINDING 304 304 ATP (BY SIMILARITY)

CC SEQUENCE 520 AA: 58403 MW: 60P2704F7A52C53E CRC64

Query Match: 76.6% Score 49: DB 1: Length 520:

Best Local Similarity 44.4% Pred. No. 3.94e+00:

Matches 4: Conservative 1: Mismatches 1: Indels 0: Gaps 0:

DB 355 FVRRVAVVV 363

QY 1 FIKRVSNVI 9

RESULT 13

ID YE06_SCHPO STANDARD: M1: 631 AA

AC Q13603

DI 15-JUL-1998 (Rel. 36, Created)

DI 15-JUL-1998 (Rel. 36, Last sequence update)

DI 15-JUL-1998 (Rel. 36, Last annotation update)

DE HYPOTHEICAL 68.8 KD PROTEIN C17H9.06C IN CHROMOSOME 1

GN SPAC17H9.06C

OS Schizosaccharomyces pombe (Fission yeast)

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales

CC Schizosaccharomycetaceae; Schizosaccharomycetes

RN

RP SEQUENCE FROM N.A

RC STRAIN:972

RA Skelton J., Chubb M., Walker P.G., Farnham M.A., Ward V.

R1 Submitted (Aug-1997) to the EMBL/GenBank/CCNU databases

CC

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL institution

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to license@isb-sib.ch).

CC

CC EMBL: Z8597; CAB1215.1

CC

CC Hypothetical protein.

KW DOMAIN 70 73 POLY-SER.

FT DOMAIN 266 269 POLY-SER.

SQ SEQUENCE 601 AA: 68796 MW: 4694F9296B109D01 CRC64:

Query Match: 76.6% Score 49: DB 1: Length 601:

Best Local Similarity 66.7% Pred. No. 3.94e+00:

Matches 6: Conservative 2: Mismatches 1: Indels 0: Gaps 0:

DB 21 FVRRVAVVV 29

QY 1 FIKRVSNVI 9

CC 11- SUBUNIT: MONOMER (BY SIMILARITY).

CC 11- SUBCELLULAR LOCATION: CYTOPLASMIC.

CC 11- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY

CC STRONG, TO METHIONYL-TRNA SYNTHETASE.

CC

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL institution

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to license@isb-sib.ch).

CC

CC EMBL: U57526; AAR4830.1

CC TIGR: M30831

CC PFAM: PF01426; tRNA-synthetase; 1

CC Aminoacyl-TRNA synthetase: Protein biosynthesis: class: ATP-binding

CC SIMILAR 3: 41

CC SIMILAR 30: 305 "HIGH" REGION

CC BINDING 304 304 ATP (BY SIMILARITY)

CC SEQUENCE 432 AA: 48835 MW: 0FB95DC18D2B41F3 CRC64:

Query Match: 76.6% Score 49: DB 1: Length 432:

Best Local Similarity 55.6% Pred. No. 3.94e+00:

Matches 5: Conservative 1: Mismatches 1: Indels 0: Gaps 0:

DB 418 FLKRVVIT 426

QY 1 FIKRVSNVI 9

RESULT 12

ID SVC12EPA STANDARD: M1: 432 AA

AC G81129

DI 15-DEC-1998 (Rel. 37, Created)

DI 15-DEC-1998 (Rel. 37, Last sequence update)

DI 15-DEC-1998 (Rel. 37, Last annotation update)

DE CYSTEINYL-TRNA SYNTHETASE (EC 6.1.1.16) [CYS-EINE--TRNA LIGASE]

GN CYSYS

OS Treponema pallidum

OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema

RN

RP SEQUENCE FROM N.A

RC STRAIN:KCHOLS

RA Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,

RA Dodson R., Quinn M., Hickey E.K., Clayton R., Ketchum K.A.,

RA Sodergren E., Hardam J.M., McLeod M.P., Salzberg S., Peterson J.,

RA Khalik H., Richardson D., Howell J.K., Chidambaram M., Utterback L.,

RA McDonald P., Artlich P., Bowman G., Cotton M.D., Firth C., Garland S.,

RA Hatch B., Horst K., Roberts K., Sandusky M., Sandusky M., Smith H.O.,

RA Venter J.C.

R1 *Complete genome sequence of Treponema pallidum, the syphilis

R1 spirochete.

R1 Science 281:375-388(1998)

CC 11- CATALYTIC ACTIVITY: ATP - L-CYSTEINE - TRNA(CYS) - AMP +

CC PYROPHOSPHATE + L-CYSTEINYL-TRNA(CYS).

```

RESULT 14
ID UBIA_POST STANDARD: PRT: 288 AA.
AC 052166:
DI 15-JUL-1999 (rel. 38, Created)
DI 15-JUL-1999 (rel. 38, Last sequence update)
DI 15-JUL-1999 (rel. 38, Last annotation update)
DE 4-HYDROXYBENZATE OCTAPENYLTRANSFERASE (EC 2.5.1.-) (4-RB
DE POLYPHENYLTRANSFERASE).
GN UBIA OR AARE:
OS Providencia stuartii.
OC Bacteria: Proteobacteria: gamma subdivision: Enterobacteriaceae:
OC Providencia.
RN 1:
RP SEQUENCE FROM N.A.
RC STRAIN=PR50:
RX PARADISE 98218723:
RA "Mutations in aarB, the ubia homolog of Providencia stuartii, result
RA in high-level aminoglycoside resistance and reduced expression of the
RA chromosomal aminoglycoside 2'-N-acetyltransferase."
RL Antimicrob. Agents Chemother. 42:959-962(1998).
RC CATALYTIC ACTIVITY: 4-HYDROXYBENZATE - FARNESYLARNESYLGERANIOL
CC - FUNCTION: SYNTHESIS OF 3-OCTAPENYL-4-HYDROXYBENZATE.
CC - CATALYTIC ACTIVITY: 4-HYDROXYBENZATE - FARNESYLARNESYLGERANIOL
CC - 3-OCTAPENYL-4-HYDROXYBENZATE.
CC - COFACTOR: REQUIRES MAGNESIUM (BY SIMILARITY).
CC - PATHWAY: SECOND STEP IN UBIQUINONE BIOSYNTHESIS.
CC - SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN, INNER MEMBRANE
CC (BY SIMILARITY).
CC - SIMILARITY: BELONGS TO THE UBIA PRENYLTRANSFERASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL: AF036903; AAB8868.1;
CC PROSITE: PSC0943; UBIA: 1.
CC PFAM: PF01040; CUX10.ctab_cyce; 1.
CC UBIQUINONE BIOSYNTHESIS: Transmembrane; Inner membrane;
CC KW TRANSMEM. 38 58 POTENTIAL.
CC FT TRANSMEM 108 128 POTENTIAL.
CC FT TRANSMEM 141 161 POTENTIAL.
CC FT TRANSMEM 163 183 POTENTIAL.
CC FT TRANSMEM 213 233 POTENTIAL.
CC FT TRANSMEM 236 258 POTENTIAL.
CC FT TRANSMEM 268 288 POTENTIAL.
CC SEQUENCE 288 AA: 32248 MW: 28819704.18FA403 CRC64;
Query Match 75.0% Score 48: DB 1: Length 288;
Best Local Similarity 55.6% Pred. No. 6:62e-00;
Matches 5: Conservative 4: Mismatches 0: Indels 0: Gaps 0;
Db 134 FIKRVSHLL 142
QY 1 FIKRVSNVI 9
RESULT 15
ID DYHC_DICD1 STANDARD: PRT: 4725 AA.
AC P34036:
DI 01-FEB-1994 (rel. 28, Created)
DI 01-FEB-1994 (rel. 28, Last sequence update)
DI 15-FEB-2000 (rel. 39, Last annotation update)
DE DYNEIN HEAVY CHAIN, CYTOSOLIC (DYHC).
GN DYHC.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota: Dictyostelidia: Dictyostelium.
RN 1:
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=AX3:
RX MEDLINE: 93107159
RA Koonee M.P., Grissom P.M., McIntosh J.R.:
RA "Dynein from Dictyostelium: primary structure comparisons between a
RA cytoplasmic motor enzyme and flagellar dynein."
RL J. Cell Biol. 119:1597-1604(1992)
RN 2:
RP CHARACTERIZATION.
RX MEDLINE: 95107019
RA Koonee M.P., Grissom P.M., Lyon M., Pope T., McIntosh J.R.:
RA Molecular characterization of a cytoplasmic dynein from
RA Dictyostelium."
RL J. Eukaryot. Microbiol. 41:645-651(1994).
CC - FUNCTION: DYNEIN HAS ATPASE ACTIVITY. CYTOPLASMIC DYNEIN ACTS AS A
CC MOTOR FOR THE INTRACELLULAR REORGANIZATION OF VESICLES AND
CC ORGANELLES ALONG MICROTUBULES.
CC - SUBUNIT: CONSISTS OF AT LEAST TWO HEAVY CHAINS AND A NUMBER OF
CC INTERMEDIATE AND LIGHT CHAINS.
CC - SUBCELLULAR LOCATION: CYTOPLASMIC.
CC - SIMILARITY: BELONGS TO THE DYNEIN HEAVY CHAIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL: Z55124; CAA78827.1;
CC PIR: S28504; S28504.
CC PIR: A44357; A44357.
CC DICTYDB: DDC1045; DHCA.
CC Motor protein: Microtubules; Dynein; ATP-binding; Coiled coil
CC DOMAIN 864 897
CC FT DOMAIN 1199 1219 COILED COIL (POTENTIAL)
CC FT DOMAIN 1338 1367 COILED COIL (POTENTIAL)
CC FT DOMAIN 1420 1436 COILED COIL (POTENTIAL)
CC FT DOMAIN 1656 1684 COILED COIL (POTENTIAL)
CC FT DOMAIN 2068 2096 MICROTUBULE-BINDING (POTENTIAL)
CC FT DOMAIN 2226 2248 COILED COIL (POTENTIAL)
CC FT DOMAIN 2437 2457 COILED COIL (POTENTIAL)
CC FT DOMAIN 3266 3344 COILED COIL (POTENTIAL)
CC FT DOMAIN 3478 3580 COILED COIL (POTENTIAL)
CC FT DOMAIN 3849 3976 ATP (POTENTIAL)
CC FT NP_BIND 2271 2278 ATP (POTENTIAL)
CC FT NP_BIND 2669 2676 ATP (POTENTIAL)
CC FT NP_BIND 3011 3018 ATP (POTENTIAL)
CC SEQUENCE 4725 AA: 538807 MW: 94019915D579D3A CRC64;
Query Match 75.0% Score 48: DB 1: Length 4725;
Best Local Similarity 55.6% Pred. No. 6:62e-00;
Matches 5: Conservative 4: Mismatches 0: Indels 0: Gaps 0;
Db 599 YMKRVANVL 707
QY 1 FIKRVSNVI 9

```

Search completed: Mon Jun 19 16:05:43 2000
Job time : 4 secs.

RA KALAMRAHETTI T., BULACH D.M., RAJAKUMAR K., ADLER B.:
 "Genetic Organization of the Lipopolysaccharide O-Antigen Biosynthetic
 Locus of *Leptospira borgpetersenii* Serovar Hardjovitis";
 RA Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF078135; AAB12950.1;
 KW Hypothetical protein.
 SQ SEQUENCE 360 AA: 41544 MW: 52201253 CRC32:
 Query Match 87.5%; Score 56; DB 2; Length 360;
 Best Local Similarity 77.8%; Pred. NO. 170e-01;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 DB 128 FIKRVSGVI 125
 I: I: I: I: I:
 QY 1 FIKRVSNVI 5
 RESULT 3
 ID Q9XKX3 PRELIMINARY: PRT: 511 AA.
 AC Q9XKX3
 DT 01-NOV-1999 (TRENBLrel. 12, Created)
 DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
 DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)
 DE PUTATIVE FLIPPASE.
 GN Wzx.
 OS Bacteroides fragilis.
 OC Bacteria: Cytophagaes; Bacteroidaceae; Bacteroides.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NCTC 9343.
 RX MEDLINE: 99307214.
 RA COMSTOCK L.E., CRYNE M.J., TZIANABCS A.C., PANTOST A.,
 ANDERSON A.B., KASPER D.L.:
 "Analysis of a capsular polysaccharide biosynthesis locus of
 Bacteroides fragilis";
 RI Infect. Immun. 67:3525-3532(1999).
 DR EMBL: AF048749; AAD40712.1;
 SQ SEQUENCE 511 AA: 58642 MW: D4196588 CRC32:
 Query Match 82.8%; Score 53; DB 2; Length 511;
 Best Local Similarity 56.7%; Pred. NO. 9.03e-01;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 DB 501 FIKRVSNIV 504
 I: I: I: I: I:
 QY 1 FIKRVSNVI 9
 RESULT 4
 ID Q9X273 PRELIMINARY: PRT: 417 AA.
 AC Q9X273
 DT 01-NOV-1999 (TRENBLrel. 12, Created)
 DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
 DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)
 DE ENDOLUCANASE.
 GN TMI75.
 OS Thermotoga maritima.
 OC Bacteria: Thermotogates; Thermotoga.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE: 99287316.
 RA NELSON K.E., CLAYTON R.A., GILL S.R., GWINN M.L., DODSON R.J.,
 HAFT D.H., HICKEY E.K., PETERSON J.D., NELSON W.C., KETCHUM K.A.,
 McDONALD L., UTTERBACK T.R., MALEK J.A., LINHER K.D., GARRETT M.M.,
 STEWART A.M., COTTON M.D., PRATT M.S., PHILLIPS C.A., RICHARDSON D.,
 REICHELBERG J., SUTTON G.G., FLEISCHMANN R.D., WHITE O., SALZBERG S.L.,
 SMITH H.O., VENTER J.C., FRASER C.M.:
 "Evidence for lateral gene transfer between Archaea and bacteria from
 genome sequence of *Thermotoga maritima*";
 RI Nature 398:323-324(1999).
 DR EMBL: AF048749; AAD40712.1;
 SQ SEQUENCE FROM N.A.
 RA NELSON K.E., CLAYTON R.A., GILL S.R., GWINN M.L., DODSON R.J.,

RA HAFT D.H., HICKEY E.K., PETERSON J.D., NELSON W.C., KETCHUM K.A.,
 McDONALD L., UTTERBACK T.R., MALEK J.A., LINHER K.D., GARRETT M.M.,
 STEWART A.M., COTTON M.D., PRATT M.S., PHILLIPS C.A., RICHARDSON D.,
 REICHELBERG J., SUTTON G.G., FLEISCHMANN R.D., WHITE O., SALZBERG S.L.,
 SMITH H.O., VENTER J.C., FRASER C.M.:
 "Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF001813; AAD36816.1;
 SQ SEQUENCE 317 AA: 37463 MW: 3841108 CRC32:
 Query Match 78.1%; Score 50; DB 2; Length 317;
 Best Local Similarity 56.7%; Pred. NO. 4.50e-00;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 DB 72 FFKRVDEVI 80
 I: I: I: I: I:
 QY 1 FIKRVSNVI 9
 RESULT 5
 ID OC6614 PRELIMINARY: PRT: 473 AA.
 AC OC6614
 DT 01-JUL-1997 (TRENBLrel. 04, Created)
 DT 01-JUL-1997 (TRENBLrel. 04, Last sequence update)
 DT 01-NOV-1998 (TRENBLrel. 08, Last annotation update)
 DE HYPOTHETICAL 50.5 KD PROTEIN.
 GN MICV336.18.
 OS Mycobacterium tuberculosis.
 OC Bacteria: Firmicutes; Actinobacteria; Actinobacteridae; Mycobacterium.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV.
 RX MURPHY L., HARRIS D.:
 "Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV.
 RA BARRELL S.G., RAZANDEKAR M.A., FARKHILL J.,
 Submited (MAY-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV.
 RX MEDLINE: 95181548.
 RA PHILLIPS W.J., PATEL S., FLEISCHMANN R., FARKHILL J.,
 RAZANDEKAR M.A., HENRY B., PARKER S., GILL S.R., SALZBERG S.L.,
 SMITH H.O.:
 "An integrated map of the genome of the model pathogen
 Mycobacterium tuberculosis H37Rv and its relation with Mycobacterium
 leprae";
 RI Proc. Natl. Acad. Sci. U.S.A. 93:3122-3127(1996).
 DR EMBL: Z95586; CAB09061.1;
 KW Hypothetical protein.
 SQ SEQUENCE 473 AA: 50474 MW: 1581A279 CRC32:
 Query Match 78.1%; Score 50; DB 2; Length 473;
 Best Local Similarity 56.7%; Pred. NO. 4.50e-00;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 DB 148 FFKRVSNPV 156
 I: I: I: I: I:
 QY 1 FIKRVSNVI 9
 RESULT 6
 ID PT1947 PRELIMINARY: PRT: 479 AA.
 AC PT1947
 DT 01-JAN-1998 (TRENBLrel. 05, Created)
 DT 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
 DT 01-JAN-1998 (TRENBLrel. 05, Last annotation update)
 DE HYPOTHETICAL 51.1 KD PROTEIN CY44.19.
 GN MICV441.19.
 OS Mycobacterium tuberculosis.
 OC Bacteria: Firmicutes; Actinobacteria; Actinobacteridae;

Q1 Actinocyetales: Corynebacterineae: Mycobacteriaceae: Mycobacterium.

OC [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN-H37AV:
 RC SKELTON J., BURCHER C.M., BARRELL B.G., RAJANDREAM M.A.:
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: STRONG, TO M.TUBERCULOSIS MTCV336.28.
 CC EMBL: 280225; CAB02329.1;
 CC Hypothetical protein.
 CC KX
 CC SEQUENCE 419 AA: 51099 MW: 50A262E8 CRC32:
 SL

Query Match 78.1% Score 50: DB 2: Length 479:

Best Local Similarity 55.7% Pred. No. 4: 9e+00:
 Matches 6: Conservative 2: Mismatches 1: Indels 0: Gaps 0:

DB 154 FIKRVSNPV 162

Q1 FIKRVSNVI 9

RESULT 7
 AC Q9W2TC PRELIMINARY: PRT: 190 AA.
 AC Q9W1G:

CC 01-NOV-1999 (TREMBLrel. 12, Created)
 CC 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 CC 01-NOV-1999 (TREMBLrel. 12, Last annotation update)
 CC TRANSCRIPTIONAL REGULATOR, TEIR FAMILY.
 CC TM0823.

OS Thermotoga maritima.

OC Bacteria: Thermotogales: Thermotoga.

CC [1]

RN SEQUENCE FROM N.A.

RX MEDLINE: 99287316.

RA NELSON K.E., CLAYTON R.A., GILL S.R., GWINN M.L., DOBSON R.J.,

RA HAFT D.H., HICKEY E.K., PETERSON C.D., NELSON W.C., KETCHUM K.A.,

RA MCDONALD L., GIERBACK J.R., MALEK J.A., LINHER K.L., GARRETT M.M.,

RA STEWART A.M., COTTON M.B., PRATT M.S., PHILLIPS C.A., RICHARDSON D.,

RA HEIDENBERG J., SUTTON G.G., FLEISCHMANN R.D., WHITE O., SALZBERG S.L.,

RA SMITH H.O., VENTER J.C., FRASER C.M.,

RA "Evidence for lateral gene transfer between Archaea and bacteria from

RA genome sequence of Thermotoga maritima."

RL Nature 399:324-329(1999).

RN [2]

RP SEQUENCE FROM N.A.

RA NELSON K.E., CLAYTON R.A., GILL S.R., GWINN M.L., DOBSON R.J.,

RA HAFT D.H., HICKEY E.K., PETERSON C.D., NELSON W.C., KETCHUM K.A.,

RA MCDONALD L., GIERBACK J.R., MALEK J.A., LINHER K.L., GARRETT M.M.,

RA STEWART A.M., COTTON M.B., PRATT M.S., PHILLIPS C.A., RICHARDSON D.,

RA HEIDENBERG J., SUTTON G.G., FLEISCHMANN R.D., WHITE O., SALZBERG S.L.,

RA SMITH H.O., VENTER J.C., FRASER C.M.,

RA Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.

CC EMBL: AB001749; AB03905.1;
 CC SEQUENCE 190 AA: 22492 MW: B32D72D0 CRC32:

Query Match 76.6% Score 49: DB 2: Length 190:

Best Local Similarity 55.7% Pred. No. 7: 56e+00:
 Matches 6: Conservative 2: Mismatches 1: Indels 0: Gaps 0:

DB 126 FIKRVVELI 134

Q1 FIKRVSNVI 9

RESULT 8
 AC Q37300 PRELIMINARY: PRT: 227 AA.
 AC Q37300:

CC 01-NOV-1996 (TREMBLrel. 01, Created)

CC 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

CC 01-NOV-1999 (TREMBLrel. 12, Last annotation update)

CC CYTOCHROME C OXIDASE POLYPEPTIDE 11 (BC 1.9.3.1).

CC COI1.

OS Chorthippus parallelus.

OC Mitochondrion.

CC [2]

RN SEQUENCE FROM N.A.

RP STRAIN-BRISTOL N2:

OC Eukaryota: Metazoa: Arthropoda: Tracheata: Hexapoda: Insecta:
 OC Pterygota: Neoptera: Orthopteroidea: Orthoptera: Caelifera:
 OC Acridomorpha: Acridoidea: Acrididae: Gomphocerinae: Chorthippus.

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-ESC:

RX MEDLINE: 96244738.

RA SZYMURA J.M., LUNT D.H., HEWITT G.M.

RL The sequence and structure of the meadow grasshopper (Chorthippus

RL parallelus) mitochondrial rRNA. ND2, COI, COII, ATPase and 9 tRNA

RL genes.

RL Insect Mol. Biol. 5:127-135(1996).

CC -1- FUNCTION: SUBUNIT I AND II FORM THE FUNCTIONAL CORE OF THE ENZYME

CC COMPLEX. ELECTRONS ORIGINATING IN CYTOCHROME C ARE TRANSFERRED VIA

CC HEME A AND CU(A) TO THE BENZYL CENTER FORMED BY HEME A3 AND

CC CU(B). SUBUNIT II BINDS CU(A) AND CYTOCHROME C.

CC -1- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C O(2) - 2 H(2)O + 4

CC FERROCYTOCHROME C.

CC -1- COFACTOR: COPPER A AND HEME GROUP.

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. (MITOCHONDRIA)

CC -1- SIMILARITY: TO OTHER MITOCHONDRIAL OR BACTERIAL COX2 SUBUNIT.

DR EMBL: X95574; CAA64823.1;

DR PROSITE: PS00078; COX2; 1.

DR PFAM: PF03116; COX2; 1.

KW Mitochondrion; Oxidoreductase; Copper; Transmembrane.

FT METAL 163 COPPER A (PROBABLE);

FT METAL 167 COPPER A (PROBABLE);

FT METAL 196 COPPER A (PROBABLE);

FT METAL 200 COPPER A (PROBABLE);

FT METAL 204 COPPER A (PROBABLE);

FT METAL 207 COPPER A (PROBABLE);

FT METAL 207 COPPER A (PROBABLE);

CC SEQUENCE 227 AA: 49E2C7D8 CRC32:

Query Match 76.6% Score 49: DB 8: Length 227:

Best Local Similarity 55.8% Pred. No. 7: 59e+00:

Matches 5: Conservative 4: Mismatches 0: Indels 0: Gaps 0:

DB 219 FIKWISKII 227

Q1 FIKRVSNVI 9

RESULT 9

ID Q61969 PRELIMINARY:

AC Q61969 PRT: 325 AA.

CC 01-NOV-1996 (TREMBLrel. 12, Created)

CC 01-NOV-1996 (TREMBLrel. 12, Last sequence update)

CC 01-NOV-1999 (TREMBLrel. 12, Last annotation update)

CC H05B21.2 PROTEIN.

CC H05B21.2

OS Caenorhabditis elegans.

OC Eukaryota: Metazoa: Nematoda: Secernentea: Rhabditia: Rhabditidae.

OC Rhabditina: Rhabditoidea: Rhabditidae: Peloderinae: Caenorhabditis

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2:

RX MEDLINE: 94150718.

RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,

RA BONFIELD J., BURTON C., CONNELL M., COPSEY T., COOPER J.,

RA CRAXTON M., DEAR S., DU Z., JUBIN R., FAVELLO A., FULTON L.,

RA GARDNER A., GREEN P., HARKIN T., HILLIER L., JIER M., JOHNSTON L.,

RA JONES M., KERSHAW J., KRISTEN J., LAISTER N., LATREILLE P.,

RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,

RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,

RA SHALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,

RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,

RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.,

RA "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.

RA elegans."

RL Nature 368:32-38(1994).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-BRISTOL N2:

RA MILLER N.A.
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.

RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BRISTOL N2
 RA WATERSTON R.
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF068717; AAC:7263.1;
 DR PFAM: PF01624; 7tm.5; 1;
 SQ SEQUENCE 325 AA: 37998 MW: AAE97C26 CRC32:

Query Match: 76.6%; Score 49; DB 5; Length 325;
 Best Local Similarity 77.8%; Pred. No. 7.58e+00;
 Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

DB 15 FIKRVSNV: 23
 1:|||||
 QY 1 FIKRVSNV: 9

RESULT 10
 ID 0924E5 IRELIMINARY: PRT: 326 AA.
 AC 0924E5.
 DT 01-MAY-1999 (TRENBLREL: 10; Created)
 DT 01-MAY-1999 (TRENBLREL: 10; Last sequence update)
 DT 01-MAY-1999 (TRENBLREL: 10; Last annotation update)
 DE PARA PROTEIN.
 GN PARA.
 OS Escherichia coli.
 OC Plasmid Incit: Colib-P9.
 CC Bacteria: Proteobacteria: gamma subdivision: Enterobacteriaceae:
 CC Escherichia.
 RN [1]
 RA SAMPEI G., MIZOBUCHI K.
 RI "Organization and diversification of plasmid genomes: complete
 RI nucleotide sequence of the Colid-P9 genome."
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB021078; BAA75111.1;
 DR Plasmid.
 KW SEQUENCE 326 AA: 36227 MW: 9246414 CRC32:

Query Match: 76.6%; Score 49; DB 2; Length 326;
 Best Local Similarity 66.2%; Pred. No. 7.58e+00;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 260 LIKRVNDV: 246
 1:|||||
 QY 1 FIKRVSNV: 9

RESULT 11
 ID 082690 IRELIMINARY: PRT: 451 AA.
 AC 082690.
 DT 01-NOV-1998 (TRENBLREL: 08; Created)
 DT 01-NOV-1998 (TRENBLREL: 08; Last sequence update)
 DT 01-NOV-1998 (TRENBLREL: 08; Last annotation update)
 DE MATI.
 GN MATI.
 OS Lepocinclis buetschlii.
 CC Eukaryota: Eumetazoa: Euglenida: Euclenales: Lepocinclis.
 RN [1]
 RA DOETSCH N.A., THOMPSON M.C., HALLICK R.B.
 RI "A tetracycline-encoding group III intron is conserved in deeply rooted
 RI Euglenoid species: are group III introns the chicken or the egg?"
 RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL: 759834; CAB1659.1;
 KW SEQUENCE 451 AA: 52368 MW: BE943A2F CRC32:

Query Match: 76.6%; Score 49; DB 10; Length 451;
 Best Local Similarity 55.8%; Pred. No. 7.58e+00;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

DB 325 FLRRVKNV: 333
 1:|||||
 QY 1 FIKRVSNV: 9

RESULT 12
 ID 073636 IRELIMINARY: PRT: 868 AA.
 AC 073636.
 DT 01-AUG-1998 (TRENBLREL: 07; Created)
 DT 01-AUG-1998 (TRENBLREL: 07; Last sequence update)
 DT 01-NOV-1999 (TRENBLREL: 12; Last annotation update)
 DE PHEROMONE RECEPTOR.
 GN CA02.1.
 OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
 CC Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Actinopterygii:
 CC Neopterygii: Teleostei: Euteleostei: Acanthopterygii: Perciformes:
 CC Tetraodontiformes: Tetraodontidae: Tetraodontidae: Fugu
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE 98226788.
 RA NAITO T., SAITO Y., YAMAMOTO T., NAKAKI Y., TOMURA K., HATARA M.,
 RA NAKANISHI S., BRENNER S.
 RI "Putative pheromone receptors related to the Ca2+-sensitive receptor 10
 RI Fugu."
 RL Proc. Natl. Acad. Sci. U.S.A. 95:578-581(1998).
 DR EMBL: ABC08858; BAA26123.1;
 DR PFAM: PF00003; 7tm.3; 1;
 DR PFAM: PF01094; ANF receptor; 1;
 DR PRINTS: PR00248; GFCRMGR.
 KW Pheromone.
 SQ SEQUENCE 868 AA: 95531 MW: 9A037662 CRC32:

Query Match: 76.6%; Score 49; DB 13; Length 868;
 Best Local Similarity 75.0%; Pred. No. 7.58e+00;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

DB 262 IKRVADV: 289
 1:|||||
 QY 2 IKRVSNV: 9

RESULT 13
 ID 09X596 IRELIMINARY: PRT: 251 AA.
 AC 09X596.
 DT 01-NOV-1999 (TRENBLREL: 12; Created)
 DT 01-NOV-1999 (TRENBLREL: 12; Last sequence update)
 DT 01-NOV-1999 (TRENBLREL: 12; Last annotation update)
 DE NIDC ADOLASE.
 GN NIDC.
 OS Rhodococcus sp. 124
 CC Bacteria: Firmicutes: Actinobacteria: Actinobacteridae:
 CC Actinomycetales: Corynebacteriales: Nocardiaceae: Rhodococcus
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=124.
 RA STEPHANOPOULOS G., SINSKEY A.J.,
 RA STEPHANOPOULOS G., LANKENAU E., LESSARD P.A.
 RI "Isolation and characterization of Indene Bioconversion Genes from
 RI Rhodococcus strain 124."
 RL Appl. Microbiol. Biotechnol. 0:0-C(1999).
 DR EMBL: AF121905; AAD25398.1;
 SQ SEQUENCE 251 AA: 27374 MW: 22E5ACD5 CRC32:

Query Match: 75.0%; Score 48; DB 2; Length 251;
 Best Local Similarity 55.6%; Pred. No. 1.27e+01;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

DB 74 FTRRVSEV: 82
 1:|||||
 QY 1 FIKRVSNV: 9

```

RESULT 14
ID Q92N37 PRELIMINARY: PRI: 367 AA.
AC Q92N37
DT 01-MAY-1999 (IREMBLrel. 10, Created)
DI 01-MAY-1999 (IREMBLrel. 10, Last sequence update)
DI 01-NOV-1999 (IREMBLrel. 12, Last annotation update)
DE POLLEN MAJOR ALLERGEN 1-2
CS Juniperus ashei (dark white cedar).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC euphyllophytes; Spermatophyta; Coniferopsida; Coniferales;
OC Taxodiaceae; Juniperus.
RN 11
RP SEQUENCE FROM N.A.
RA MIDORO-HORIUCHI T.M., GOLDRUM R.M., KUROSKY A., WOOD T.G.,
RA BROOKS E.G.
RI "Molecular cloning of mountain cedar (Juniperus ashei) pollen major
RI allergen, Jun. a 1";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF106663; AAD03659.1;
DR EMBL: AF106662; AAD03658.1;
DR MENDEL: 36544; Juncas:1088;36544.
DR MENDEL: 36545; Juncas:1088;36545.
SU SEQUENCE 367 AA; 39824 MW; 4C2D58630 CRC32;

Query Match 75.0%; Score 48; DB 10; Length 367;
Best Local Similarity 55.6%; Pred. No. 1.27e+01;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

DR 130 FMRKVSHT 138
CV : FIKRVSNVI 9

RESULT 15
ID Q86020 PRELIMINARY: PRI: 475 AA.
AC Q86020
DT 01-NOV-1998 (IREMBLrel. 08, Created)
DI 01-NOV-1998 (IREMBLrel. 08, Last sequence update)
DI 01-NOV-1999 (IREMBLrel. 12, Last annotation update)
DE PYRUVATE KINASE (EC 2.7.1.40) (PHOSPHOENOLPYRUVATE KINASE)
DE (PHOSPHOENOL TRANSPHOSPHORYLASE).
GN PYK.
OS Zymomonas mobilis.
OC Bacteria; Proteobacteria; alpha subdivision; Zymomonas group;
OC Zymomonas.
RN 11
RP SEQUENCE FROM N.A.
RA STEINER P., FUSSENNEGER M., BATTLE J.E., SAUER U.
RI "Cloning and expression of the Zymomonas mobilis pyruvate kinase gene
RI in Escherichia coli";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC 1- CATALYTIC ACTIVITY: ATP + PYRUVATE -> ADP + PHOSPHOENOLPYRUVATE.
CP EMBL: AF079586; AAC29104.1;
CP WSSP: P1A178; LPX.
DR ZFAM: PF00224; PK.
DR PRINTS: PR01050; PYRUVTKASE.
KW Pyruvate; Transferease.
SQ SEQUENCE 475 AA; 51445 MW; 28F3CA3B CRC32;

Query Match 75.0%; Score 48; DB 2; Length 475;
Best Local Similarity 55.6%; Pred. No. 1.27e+01;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

DR 199 FVORVEDVI 207
CV : FIKRVSNVI 9

```

Search completed: Mon Jun 19 16:06:09 2000
 Job time : 9 secs.

PI 03-SEP-1993: F63703.
 PE 10-SEP-1990: FR-01186.
 PA (INSP) INST PASTEUR.
 P1 Mazodier P, Guglielmi G.
 DE WPI: 92-114358/14.
 UK NSDB: Q22483.
 V1 Recombinant DNA conig. heat inducible promoter and heterologous
 V2 gene - also vectors, transformed cells and new heat shock
 V3 proteins of Streptococcus albus.
 PS Disclosure: Fig 5: 50pp; French.
 P3 The sequence is that of the GroES protein which is encoded by the
 CC structural gene groES. See also Q22477-Q22486.
 SC Sequence 102 AA.

Query Match 90.48; Score 103; DB 1; Length 102;
 Best Local Similarity 76.58; Pred. No. 4.15e-03;
 Matches 13; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

DB 95 EYLVLARSREVLAIIEK 102
 QY 1 EYLVLARSREVLAVSK 17
 ||| ||||| |||

RESULT 3
 ID R05700 standard; protein: 99 AA.

AC R05700;
 DT 16-AUG-1990 (first entry)
 DE MPB-57 protein.
 KW BCG; tuberculosis; MPB-57; ELISA; ds.
 US Mycobacterium bovis BCG.
 PN J0205496-A.
 DT 22-FEB-1990.
 P1 18-AUG-1988: 205444.
 P2 18-AUG-1988: JP-205444.
 P3 (AJIN) Ajinomoto KK.
 P4 WPI: 92-103125/14.
 P5 N-PSDR: Q23735.
 P6 BCG bacteria derived MPB-57 protein -
 P7 used esp. in diagnosis of tuberculosis.
 P8 Claim 1: Fig 2: 10pp; Japanese.
 P9 Large quantities of MP-57 protein can be derived from a cultured
 CC transformant and used in diagnosis of tuberculosis by ELISA.
 SC Sequence 99 AA.

Query Match 89.38; Score 73; DB 1; Length 99;
 Best Local Similarity 84.68; Pred. No. 1.20e-03;
 Matches 21; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

DB 93 EYLVLARSREVLAVG 95
 QY 1 EYLVLARSREVLAV 15
 ||||| |||

RESULT 4
 ID R67380 standard; protein: 94 AA.

AC R67380;
 DT 22-JUN-1995 (first entry)
 DE GroES-like protein.
 KW GroES-like protein; vaccine; diagnostic; heat shock protein; HSP;
 US GroES-like protein; Helicobacter fetus.
 PN W09426901-A.
 DT 24-NOV-1994.
 P1 19-MAY-1994: E01625.
 P2 19-MAY-1993: EP-401309.
 P3 19-NOV-1993: WO-E03259.
 P4 (INRM) INST NAT SANTE & RECH MEDICALE.
 P5 (INSP) INST PASTEUR.
 P6 Ferrero R, Labigne A, Suerbaum S, Thiberge J;
 P7 DNA from Helicobacter pylori and Helicobacter fetus - used to
 P8 develop prods. for detection, treatment and prevention of
 P9 Helicobacter infection

PS Disclosure: Fig. 7B(1-11): 168pp; English.
 CC The sequence of the Helicobacter pylori heat shock protein H
 CC (given in R67373) was compared to that of other GroES-like
 CC proteins from Legionella pneumophila (R67387), Escherichia coli
 CC (R67390), Clostridium perfringens (R67389), Mycobacterium leprae
 CC (R67386) and thermophilic bacterium (R67388), and regions
 CC of homology were identified.
 SC Sequence 94 AA.

Query Match 61.44; Score 70; DB 1; Length 94;
 Best Local Similarity 57.14; Pred. No. 9.21e-00;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

DB 80 EYLVLARSREVLAVI 93
 QY 2 EYLVLARSREVLAV 15
 ||||| |||||

RESULT 5

ID R67389 standard; protein: 94 AA.
 AC R67389;
 DT 22-JUN-1995 (first entry)
 DE C. perfringens GroES-like protein.
 KW GroES-like protein; vaccine; diagnostic; heat shock protein; HSP;
 US GroES-like protein; Helicobacter fetus.
 PN W09426901-A.
 DT 24-NOV-1994.
 P1 19-MAY-1994: E01625.
 P2 19-MAY-1993: EP-401309.
 P3 19-NOV-1993: WO-E03259.
 P4 (INRM) INST NAT SANTE & RECH MEDICALE.
 P5 (INSP) INST PASTEUR.
 P6 Ferrero R, Labigne A, Suerbaum S, Thiberge J;
 P7 WPI: 95-006797/01.
 P8 DNA from Helicobacter pylori and Helicobacter fetus - used to
 P9 develop prods. for detection, treatment and prevention of
 P0 Helicobacter infection
 P1 Disclosure: Fig. 7B(1-11): 168pp; English.
 P2 The sequence of the Helicobacter pylori heat shock protein H
 CC (given in R67373) was compared to that of other GroES-like
 CC proteins from Legionella pneumophila (R67387), Escherichia coli
 CC (R67390), Clostridium perfringens (R67389), Mycobacterium leprae
 CC (R67386) and thermophilic bacterium (R67388), and regions
 CC of homology were identified.
 SC Sequence 94 AA.

Query Match 57.08; Score 65; DB 1; Length 94;
 Best Local Similarity 53.08; Pred. No. 2.07e-00;
 Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

DB 79 EYLVLARSREVLAVI 93
 QY 1 EYLVLARSREVLAV 15
 ||||| |||||

RESULT 6

ID R67390 standard; protein: 97 AA.
 AC R67390;
 DT 22-JUN-1995 (first entry)
 DE E. coli GroES-like protein.
 KW GroES-like protein; vaccine; diagnostic; heat shock protein; HSP;
 US GroES-like protein; Helicobacter fetus.
 PN W09426901-A.
 DT 24-NOV-1994.
 P1 19-MAY-1994: E01625.
 P2 19-MAY-1993: EP-401309.
 P3 19-NOV-1993: WO-E03259.
 P4 (INRM) INST NAT SANTE & RECH MEDICALE.
 P5 (INSP) INST PASTEUR.
 P6 Ferrero R, Labigne A, Suerbaum S, Thiberge J;
 P7 WPI: 95-006797/01.

PF DNA from *Helicobacter pylori* and *Helicobacter felis* - used to
 PF develop prods. for detection, treatment and prevention of
 PF *Helicobacter* infection
 PS Disclosure: Fig. 7B(1-11): 168pp; English.
 CC The sequence of the *Helicobacter pylori* heat shock protein B
 CC (given in R67373) was compared to that of other *Gr-ES*-like
 CC proteins from *Legionella pneumophila* (R67387), *Escherichia coli*:
 CC (R67390), *Clostridium perfringens* (R67389), *Mycobacterium leprae*
 CC (R67386) and thermophilic bacterium (R67388), and regions
 CC of homology were identified.
 SQ Sequence 97 AA;

Query Match 57.0%; Score 65; DB 1; Length 136;
 Best Local Similarity 60.0%; Pred. No. 2,778-01;

Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

DB 61 EYVLMSESDILAIIV 95

|||||

QY 1 EYVILSARDVLAVV 15

RESULT 7
 ID W98289 standard; Protein: 136 AA.

AC W98289;

DE 31-MAR-1999 (first entry)

DE H. pylori GPC 1543 protein.

KW GPC protein; *Helicobacter* infection; gastroduodenal disease; gastritis;

KW peptic ulcer disease.

OS *Helicobacter pylori*.

PN WC9823478-A1.

PS 08-OCT-1998.

PF 01-APR-1998; J06371.

PK 29-JUL-1997; US-502615.

PK 01-APR-1997; US-833457.

PK 24-JUN-1997; US-881227.

PA (HUMAN) HUMAN GENOME SCI. INC.

PA (INNR) MERIEUX ORAVAX PASTEUR MERIEUX SERUMS.

PI Al. Garza, A. Kieckhefer, H. Miller, C. Oomen RP, Tomb J;

PI WPI: 98-542293/45.

DR N-PSDB: X14008.

PF New isolated *Helicobacter* polynucleotides - used to develop products

PF for the diagnosis, prevention and treatment of *Helicobacter*

PF infections and gastrointestinal diseases

PS Claim 8, Page 348-349; 2054pp; English.

CC This sequence represents a *Helicobacter pylori* GPC protein of the

CC invention. The polypeptides can be used for preventing or treating

CC *Helicobacter* infections, and gastroduodenal diseases associated with

CC these infections, including acute, chronic, and atrophic gastritis, and

CC peptic ulcer diseases, e.g. gastric and duodenal ulcers. They can also be

CC used for the production of antibodies. The products can also be used for

CC detection and diagnosis.

SQ Sequence 136 AA;

Query Match 57.0%; Score 65; DB 1; Length 136;

Best Local Similarity 50.0%; Pred. No. 2,778-01;

Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

DE 8 EDVLMMSREVSAPVS 23

|||||

QY 1 EYVILSARDVLAVV 15

RESULT 8
 ID R67373 standard; Protein: 118 AA.

AC R67373;

DE 22-JUN-1995 (first entry)

DE Heat shock protein HSPB

KW Urease; urea gene; urea gene; immunogen; vaccine; diagnostic;

KW *Helicobacter pylori*; heat shock protein; HSP; chaperonin;

KW pili2205.

OS *Helicobacter felis* ATCC 49179.

PN WO9424901-A.

PS 24-NOV-1994.

PF 19-MAY-1994; E01625.
 PR 19-MAY-1993; EP-401309.
 PR 19-NOV-1993; WO-E03259.
 PA (INNR) INST NAT SANTE & RECH MEDICALE.
 PA (INSP) INST PASTEUR.
 PI Ferrero R, Labigne A, Suerbaum S, Thiberge J;
 DR WPI: 95-006797/01.
 DR N-PSDB: 075321.
 PT DNA from *Helicobacter pylori* and *Helicobacter felis* - used to
 PT develop prods. for detection, treatment and prevention of
 PT *Helicobacter* infection
 PS Disclosure: Fig. 6; 168pp; English.
 CC Vaccine compositions include the A and B subunits (given in R67373)
 CC of *H. felis* urease encoded by the ureA/ureB gene (475319) region
 CC of the urease gene cluster of pili205 (CNOM 1-1355), as well as the
 CC heat shock proteins HSPA (R67374) and HSPB (R67373) encoded by the
 CC urease-associated HSP gene cluster region (Q75321) of pili689 (CNOM
 CC 1-1356). Recombinant products are expressed in *Escherichia coli*.
 SQ Sequence 118 AA;

Query Match 53.5%; Score 61; DB 1; Length 118;

Best Local Similarity 40.0%; Pred. No. 6,578-01;

Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

DB 76 EYVWLEEDILGIVG 90

|||||

QY 2 EYVILSARDVLAVV 16

RESULT 9

ID R74338 standard; Protein: 118 AA.

AC R74338;

DE 04-NOV-1995 (first entry)

DE *Helicobacter pylori* heat shock protein HSPA.

KW Heat shock protein; HSPA; HSPB; urease; UreA; UreB; UreH;

KW vaccine; *Helicobacter* infection.

OS *Helicobacter pylori*.

PI Key Location/Qualifiers

PI MISC_difference 92..118

PI /note: "p1: red vaccine component"

PN WC9514093-A.

PS 26-MAY-1995.

PR 19-NOV-1993; E03259.

PA (INNR) INST NAT SANTE & RECH MEDICALE.

PA (INSP) INST PASTEUR.

PI Ferrero R, Labigne A, Suerbaum S;

DR WPI: 95-200383/25.

DR N-PSDB: Q90181.

PT Immunogenic composition against *Helicobacter* infection - also

PT gene fragment(s) and protein(s) from *Helicobacter* urease gene

PI cluster and heat shock proteins(s).

PS Claim 17 + 20; Fig. 6; 128pp; English.

CC The HSPA heat shock protein is a component of a novel immunogenic

CC composition capable of inducing protective antibodies against

CC *Helicobacter* infection. The composition may include the HSPB

CC protein (R74339), a urease UreA or UreB protein (R74336, R74337)

CC or the UreC protein (R74340). The composition is used to prepare

CC a vaccine for humans or animals, especially against *H. pylori* and

CC *H. felis*. Antibodies against the proteins may be used for

CC treating *Helicobacter* infection, and primers/probes to the DNA

CC sequence may be used for detection of *Helicobacter* infection.

SQ Sequence 118 AA;

Query Match: 53.5%; Score 61; DB 1; Length 118;

Best Local Similarity 40.0%; Pred. No. 6,578-01;

Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

DB 76 EYVWLEEDILGIVG 90

|||||

QY 2 EYVILSARDVLAVV 16

```

RESULT 10
AC W05731 standard; Protein: 118 AA.
DE H. pylori heat shock protein HspA.
KW Heat shock protein; HspA; Hsp3; Chaperonin; urease; Urea; UreB;
KW vaccine; immunogen; antigen; antibody; chronic gastritis; ulcer;
KW p11589.
OS Helicobacter pylori strain 85P.
FH Key location/Qualifiers
FT domain 92..118
PI
PI Note= "this region of HspA resembles a metal
PI binding domain and is a preferred fragment
PI for use in compns. of the invention
PI (claim 3)."
PI
PI W05634624-A1.
PI
PI 07-NOV-1996.
PI 02-MAY-1996; E01834.
PI 19-MAY-1995; US-447177.
PI 19-MAY-1995; US-432697.
PI (INRM ) INST NAT SANTE & RECH MEDICALE.
PI (INSE ) INST PASTEUR.
PI Ferrero RL; Labiane A, Suerbaum S, Thiberge J;
PI WPI: 98-505608/50.
PI N-PSDB: T49681.
PI New immunogenic compo. contg. UreB and HspA antigens of
PI Helicobacter - for treatment and prevention of esp. H pylori
PI infection, also new antibodies specific for these antigens.
PI Example 3; Page 112; 184pp; English.
PI Heat shock proteins HspA (W06731) and HspB (W06732) are products
PI of a Helicobacter pylori gene cluster (see also T45681) in plasmid
PI p11589 (CNCM 1-1156). Recombinant HspA and HspB can be produced
PI in transformed host cells for use (esp. HspA) with Helicobacter
PI urease subunits (see also W06729-30) in novel immunogenic compns.
PI These compns. can provoke a mucosal response and are useful as
PI vaccines to protect humans or other animals (esp. cats and dogs)
PI against Helicobacter infection or for treatment of such infections.
PI The C-terminal fragment (see also W06725) of HspA is a preferred
PI component of such compns.
PI Sequence 118 AA.
PI
PI Query Match 53.5% Score 61; DB 1; Length 118;
PI Best Local Similarity 40.0% Pred. No. 6,576-01;
PI Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;
PI
PI 76 EYVLEEDLGLGVG 90
PI 111111111111
PI 2 EYVLSARDVLAVV 16
PI
PI RESULT 11
PI W22256 standard; Protein: 106 AA.
PI
PI 02-DEC-1998 (first entry)
PI Human receptor type tyrosine kinase-like protein.
PI Tyrosine kinase; receptor; Eph family; human; HEP receptor;
PI modification; amplification.
PI Homo sapiens.
PI W05745586-A1.
PI 11-DEC-1997.
PI 04-JUN-1997; J01887.
PI 04-JUN-1996; JP-141849.
PI (SHIC ) SHIONOGI & CO LTD.
PI Matsui T;
PI WPI: 98-042121/04.
PI N-PSDB: V5277.
PI Eph family receptor type tyrosine kinase-like proteins and their DNA
PI - for PCR probes in studying HEP receptors
PI Claim 4; Page 24-28; 44pp; Japanese.
PI This sequence represents a novel human Eph family-like receptor type
PI tyrosine kinase-like protein. The encoding DNA may be used as PCR
PI probes in studying HEP receptors. The proteins show dominant negative
PI (variant being dominant) or signal modification and/or amplification.

```

```

SQ Sequence 1006 AA;
Query Match 52.6% Score 60; DB 1; Length 1006;
Best Local Similarity 51.5% Pred. No. 8,136-01;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
DB 591 FLLAAITVLAVV 603
QY 3 YILSARDVLAVV 15
PI
PI RESULT 12
PI W70525 standard; Protein: 1006 AA.
PI AC W70525;
PI 02-FEB-1999 (first entry)
PI Human thymus receptor tyrosine kinase (TRTK).
PI Thymus receptor; tyrosine kinase; TRTK; human; colon carcinoma;
PI testicular carcinoma; pancreatic carcinoma; lung adenocarcinoma;
PI breast carcinoma; hepatocellular carcinoma; cancer;
PI cell proliferation; diagnosis; therapy; prognosis.
PI OS Homo sapiens.
PI FH Key location/Qualifiers
PI FT Peptide 1..16
PI FT /label= Sig_peptide
PI FT Protein 27..1006
PI FT /label= Mat_protein
PI FT Domain 1..578
PI FT /label= Extracellular
PI FT Domain 579..607
PI FT /label= Transmembrane
PI FT Domain 608..1006
PI FT /label= Intracellular
PI W05844111-A1.
PI 08-OCT-1998.
PI 27-MAR-1998; U56021.
PI 28-MAR-1997; US-042856.
PI (HUMA-) HUMAN GENOME SCI INC.
PI Ruben SM, Soppet DR;
PI WPI: 98-542702/46.
PI N-PSDB: V33697.
PI New isolated thymus receptor tyrosine kinase - is used to develop
PI products for the diagnosis and treatment of cancers and other
PI diseases involving aberrant cell proliferation
PI Claim 16; Fig 1A-D; 100pp; English.
PI This is the amino acid sequence of 1 of 2 (see also W70526)
PI putative novel human thymus receptor tyrosine kinases (TRTKs) as
PI deduced from the nucleotide sequence of a cDNA clone (see V33697)
PI obtained from a human foetal cell cDNA library. It has a deduced
PI mol.wt. of 109.3 kDa. The novel TRTKs share sequence homology with
PI cck9. They are thus members of the Eph family of RTKs and
PI potentially involved in diseases resulting in alterations in normal
PI cell proliferation, such as cancer. They are believed to be
PI involved in haematopoietic development. The invention provides
PI TRTK polynucleotides and polypeptides (including the mature protein,
PI soluble polypeptides, the extracellular, transmembrane and
PI intracellular domains, and epitope-bearing portions of such
PI polypeptides), vectors, host cells and recombinant methods of
PI producing these. Also provided are diagnostic methods for detecting
PI disease states associated with the aberrant expression of TRTK and
PI therapeutic methods for treating such disease states, especially
PI cancer. TRTK is believed to be involved in a number of disease
PI states such as cancers, e.g. testicular carcinoma, pancreatic
PI carcinoma, colon carcinoma, lung adenocarcinoma, mammary carcinoma
PI and hepatocellular carcinoma.
PI Sequence 1006 AA;
PI
PI Query Match 52.6% Score 60; DB 1; Length 1006;
PI Best Local Similarity 61.5% Pred. No. 8,136-01;
PI Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
PI
PI DB 591 FLLAAITVLAVV 603
PI 111111111111
PI 3 YILSARDVLAVV 15
QY

```

```

RESULT 13
ID W0526 standard: Protein: 1021 AA.
AC W0526:
DE 02-FEB-1999 (first entry)
DD Human thymus receptor tyrosine kinase (TRTK).
KW Thymus receptor tyrosine kinase; TRTK; human; colon carcinoma;
KW testicular carcinoma; pancreatic carcinoma; lung adenocarcinoma;
KW breast carcinoma; hepatocellular carcinoma; cancer;
KW cell proliferation; diagnosis; therapy; prognosis
DS Homo sapiens.
FH Key Location/Qualifiers
FT Peptide 1..14
FT Protein /label= S1c_peptide
FT Domain /label= Mat_protein
FT Domain 1..1594
FT Domain /label= Extracellular
FT Domain 595..622
FT Domain /label= Transmembrane
FT Domain 623..1021
FT /label= Intracellular
PR W09844111.AL
PR 08-OCT-1999:
PR 27-MAR-1999: U06321.
PR 28-MAR-1997: US-042856.
PR (HUMA-) HUMAN GENE SC: INC.
PR Kuber SW, Secret DR:
PR WPI: 98-542702/45.
PR N-PSDB: V33697.
PR New isolated thymus receptor tyrosine kinase - is used to develop
PR products for the diagnosis and treatment of cancers and other
PR diseases involving aberrant cell proliferation
PR Claim 16: Fig 1A-D: 10pp: English.
PR This is the amino acid sequence of 1 of 2 (see also W0525)
PR putative novel human thymus receptor tyrosine kinases (TRTKs), as
PR deduced from the nucleotide sequence of a cDNA clone (see W33697)
PR obtained from a human foetal cell cDNA library. It has a deduced
PR mol.wt. of 111.0 kDa. The novel TRTKs share sequence homology with
PR cdk9. They are thus members of the Eph family of RTKs and
PR potentially involved in diseases resulting in alterations in normal
PR cell proliferation, such as cancer. They are believed to be
PR involved in haematopoietic development. The invention provides
PR TRTK polypeptides and polypeptides (including the mature protein,
PR soluble polypeptides, the extracellular, transmembrane and
PR intracellular domains, and epitope-bearing portions of such
PR polypeptides). vectors, host cells and recombinant methods of
PR producing these. Also provided are diagnostic methods for detecting
PR disease states associated with the aberrant expression of TRTK and
PR therapeutic methods for treating such disease states, especially
PR cancer. TRTK is believed to be involved in a number of disease
PR states such as cancers, e.g. testicular carcinoma, pancreatic
PR carcinoma, colon carcinoma, lung adenocarcinoma, mammary carcinoma
PR and hepatocellular carcinoma.
PR Sequence 1021 AA:
Query Match 52.6% Score 60: DB 1: Length 1021:
Residual Similarity 62.5% Pred.No. 8.13e+01:
Matches 8: Conservative 3: Mismatches 2: Indels 0: Gaps 0:
DE 606 FLUAAIVLAVV 518
CV 1111 1111
3 YLLSAROVAVV 15
RESULT 14
ID W74801 standard: Protein: 74 AA.
AC W74801:
DE 25-JAN-1999 (first entry)
DD Human secreted protein encoded by gene 72 clone HBIA195.
KW Human: secreted protein; testis; tumour; foetal brain tissue;
KW fusion protein; cancer; central nervous system; seizure;

```

```

KW diagnosis: neurodegenerative disease.
OS Homo sapiens.
FH Key Location/Qualifiers
FT Misc-difference 8 label= unknown
FT Misc-difference 23 label= unknown
FT Misc-difference 38 label= unknown
FT Misc-difference 43 label= unknown
FT Misc-difference 44 label= unknown
FT Misc-difference 73 label= unknown
FT W09839448.A2.
PD 11-SEP-1998: J04493.
PR 06-MAR-1999: US-061060.
PR 02-OCT-1997: US-038621.
PR 07-MAR-1997: US-040161.
PR 07-MAR-1997: US-040162.
PR 07-MAR-1997: US-040163.
PR 07-MAR-1997: US-040333.
PR 07-MAR-1997: US-040334.
PR 07-MAR-1997: US-040335.
PR 07-MAR-1997: US-040626.
PR 11-APR-1997: US-043311.
PR 11-APR-1997: US-043312.
PR 11-APR-1997: US-043313.
PR 11-APR-1997: US-043314.
PR 11-APR-1997: US-043558.
PR 11-APR-1997: US-043559.
PR 11-APR-1997: US-043576.
PR 11-APR-1997: US-043578.
PR 11-APR-1997: US-043580.
PR 11-APR-1997: US-043659.
PR 11-APR-1997: US-043670.
PR 11-APR-1997: US-043671.
PR 11-APR-1997: US-043672.
PR 11-APR-1997: US-043674.
PR 23-MAY-1997: US-047492.
PR 23-MAY-1997: US-047500.
PR 23-MAY-1997: US-047501.
PR 23-MAY-1997: US-047502.
PR 23-MAY-1997: US-047503.
PR 23-MAY-1997: US-047581.
PR 23-MAY-1997: US-047582.
PR 23-MAY-1997: US-047583.
PR 23-MAY-1997: US-047584.
PR 23-MAY-1997: US-047585.
PR 23-MAY-1997: US-047586.
PR 23-MAY-1997: US-047587.
PR 23-MAY-1997: US-047588.
PR 23-MAY-1997: US-047589.
PR 23-MAY-1997: US-047590.
PR 23-MAY-1997: US-047592.
PR 23-MAY-1997: US-047593.
PR 23-MAY-1997: US-047594.
PR 23-MAY-1997: US-047595.
PR 23-MAY-1997: US-047596.
PR 23-MAY-1997: US-047597.
PR 23-MAY-1997: US-047598.
PR 23-MAY-1997: US-047599.
PR 23-MAY-1997: US-047600.
PR 23-MAY-1997: US-047601.
PR 23-MAY-1997: US-047612.
PR 23-MAY-1997: US-047613.
PR 23-MAY-1997: US-047614.
PR 23-MAY-1997: US-047615.
PR 23-MAY-1997: US-047617.
PR 23-MAY-1997: US-047618.
PR 23-MAY-1997: US-047632.

```

23-MAY-1997; US-047633.
 06-JUN-1997; US-048954.
 06-JUN-1997; US-048974.
 13-JUN-1997; US-049610.
 08-JUL-1997; US-051926.
 16-JUL-1997; US-052874.
 18-AUG-1997; US-055724.
 22-AUG-1997; US-056630.
 22-AUG-1997; US-056641.
 22-AUG-1997; US-056652.
 22-AUG-1997; US-056656.
 22-AUG-1997; US-056657.
 22-AUG-1997; US-056662.
 22-AUG-1997; US-056664.
 22-AUG-1997; US-056845.
 22-AUG-1997; US-056845.
 22-AUG-1997; US-056862.
 22-AUG-1997; US-056864.
 22-AUG-1997; US-056864.
 22-AUG-1997; US-056872.
 22-AUG-1997; US-056874.
 22-AUG-1997; US-056875.
 22-AUG-1997; US-056876.
 22-AUG-1997; US-056877.
 22-AUG-1997; US-056878.
 22-AUG-1997; US-056879.
 22-AUG-1997; US-056880.
 22-AUG-1997; US-056881.
 22-AUG-1997; US-056882.
 22-AUG-1997; US-056884.
 22-AUG-1997; US-056885.
 22-AUG-1997; US-056887.
 22-AUG-1997; US-056887.
 22-AUG-1997; US-056888.
 22-AUG-1997; US-056889.
 22-AUG-1997; US-056892.
 22-AUG-1997; US-056893.
 22-AUG-1997; US-056894.
 22-AUG-1997; US-056903.
 22-AUG-1997; US-056908.
 22-AUG-1997; US-056909.
 22-AUG-1997; US-056910.
 22-AUG-1997; US-056911.
 05-SEP-1997; US-057050.
 05-SEP-1997; US-057059.
 05-SEP-1997; US-057061.
 12-SEP-1997; US-057761.
 (HUMAN) HUMAN GENOME SCI. INC.
 Bednarik DP, Brewer LA, Carter KC, Dean R, Ebner R, Endress GA,
 Feng P, Ferris AM, Fischer CL, Florence KA, Greene JM, Hu JS,
 Kyaw H, Laffey DW, Li Y, Moore PA, Ni J, Olsen HS, Rosen CA,
 Ruben SM, Shi Y, Sobhet DP, Young PE, Yu GL, Zeng Z,
 WPI: 98-50614/4;
 N-PSDB: V59582.
 New isolated human genes and the secreted polypeptide(s), they encode
 useful for diagnosis and treatment of e.g. cancers, neurological
 disorders, immune diseases, inflammation or blood disorders
 Claim 1: Page 582-583; 72pp: English.
 This sequence represents a secreted human protein encoded by the nucleic
 acid molecule designated Gene 72 from the human cDNA clone BBA195
 (deposited as clone ATCC 97900 and ATCC 209046).
 The gene can be used to generate fusion proteins by linking to the gene
 to a human immunoglobulin Fc portion (e.g. V59502) for increasing the
 stability of the fused protein as compared to the human protein on-y.
 The invention relates to 186 novel genes and their fragments (nucleic
 acid sequences: V59511-V59812; amino acid sequences W4731-W75026) which
 are useful for preventing, treating or ameliorating medical conditions
 e.g. by protein or gene therapy. Also, pathological conditions can be
 diagnosed by determining the amount of the new polypeptides in a sample
 or by determining the presence of mutations in the new polynucleotides.
 Specific uses are described for each of the 186 polynucleotides, based on
 which tissues they are most highly expressed in (see V59511 for described
 uses).
 Sequence 74 AA;

Query Match 51.88; Score 59; DB 1; Length 74;

Best Local Similarity 53.88; Pred. No. 1: 0le-02;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 Db 52 EYVILMNKALLT 64
 I I I I I I I I I I
 QY 1 EYVILSARDVLA 13
 RESULT 15
 ID R13334 standard: Protein: 102 A.
 AC R13334:
 DT 22-OCT-1991 (first entry)
 DE HypA protein.
 KW Antibodies; heat shock; hypersensitive; allergen; HSP40; GroES.
 OS Chlamydia psittaci GP/C.
 PN US7533317-A.
 PD 09-JUL-1991.
 PF 31-MAY-1990; 143560.
 PR 31-MAY-1990; US-531317.
 PA (USSH) NAT INST OF HEALTH.
 DR WPI: 91-245693/33.
 DR N-PSDB: Q13136.
 PI DNA encoding HypA and HypB Chlamydia proteins - used to develop
 PI Prods. for detection of and vaccines against Chlamydia infection.
 PS Disclosure; fig 5; Sipp; English.
 CC The sequence was deduced from the first of two ORFs found in clone
 CC pps57, prepd. from C. psittaci genomic DNA and contg. the Hyp
 CC operon. It is the HypA protein, of approx. 12 kD, analogous to
 CC the GroES heat shock protein of E. coli. The recombinant protein
 CC can be used to raise antisera and in the preparation of
 CC vaccines for the treatment of Chlamydial infections.
 CC See also R13335-R13337.
 SQ Sequence 102 AA;
 Query Match 50.98; Score 58; DB 1; Length 102;
 Best Local Similarity 46.78; Pr. J. No. 1: 24e-02;
 Matches 7; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
 Db 87 EYVIVQEESEVMAYL 101
 I I I I I I I I I I
 QY 1 EYVILSARDVLA 15

Search completed: Mon Jun 19 16:08:45 2000
 Job time : 8 secs.

NWESKE

(TW)

Release 3.1A John F. Collins, BioComputing Research Unit,
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution Rights by Oxford Molecular Ltd
Mfarch_pp protein - protein database search using Smith-Waterman algorithm
R.M. On: Tue Jun 20 13:35:33 2000: MasPar time 2.33 Seconds
105.377 Million cell updates/sec
Tabular output not generated.

Title: >US-09-142-524A-8
Description: (1-17) from USC9.42524A.pep
Perfect Score: 114
Sequence: 1 EYGLSAROVZAVYSK 17
Scoring table: PAM 150
Gap 15
Searched: 145341 seqs, 1443740 residues
Post-processing: Minimum Match 0%
Listing first 45 summaries
Database: a-issued
1:5A.COMS 2:5B.COMB 3:6.COMB 4:PCI.COMB 5:backfiles1
Statistics: Mean 18.743: Variance 66.315: scale 0.2H3

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query Match | Length | ID | Description | Pred. No. |
|------------|-------------|--------|------|---|-----------|
| 1 | 114 | 100.0 | 93 | US-08-467- Sequence 36, Application US/08457822 | 1.19e-04 |
| 2 | 70 | 61.4 | 94 | US-08-467- Sequence 36, Application US/08457822 | 4.52e-00 |
| 3 | 65 | 57.0 | 94 | US-08-467- Sequence 36, Application US/08457822 | 1.19e-01 |
| 4 | 55 | 57.0 | 97 | US-08-467- Sequence 36, Application US/08457822 | 1.19e-01 |
| 5 | 61 | 53.5 | 118 | US-08-467- Sequence 36, Application US/08457822 | 1.19e-01 |
| 6 | 57 | 50.0 | 344 | US-09-031- Sequence 3, Application US/08457822 | 1.19e-01 |
| 7 | 57 | 50.0 | 1612 | US-08-169- Sequence 2, Application US/08457822 | 1.19e-01 |
| 8 | 55 | 48.2 | 315 | US-08-118- Sequence 34, Application US/08457822 | 1.19e-02 |
| 9 | 55 | 48.2 | 315 | US-08-673- Sequence 11, Application US/08457822 | 1.19e-02 |
| 10 | 55 | 48.2 | 612 | US-08-673- Sequence 11, Application US/08457822 | 1.19e-02 |
| 11 | 55 | 48.2 | 973 | US-08-162- Sequence 8, Application US/08457822 | 1.19e-02 |
| 12 | 54 | 47.4 | 327 | US-08-442- Sequence 35, Application US/08457822 | 1.47e-02 |
| 13 | 54 | 47.4 | 327 | US-08-442- Sequence 35, Application US/08457822 | 1.47e-02 |
| 14 | 54 | 47.4 | 572 | US-08-745- Sequence 1, Application US/08457822 | 1.47e-02 |
| 15 | 54 | 47.4 | 997 | US-08-387- Sequence 4, Application US/08457822 | 1.47e-02 |
| 16 | 53 | 45.5 | 789 | US-08-431- Sequence 20, Application US/08457822 | 1.80e-02 |
| 17 | 53 | 45.5 | 789 | US-08-431- Sequence 20, Application US/08457822 | 1.80e-02 |
| 18 | 52 | 45.6 | 245 | US-08-825- Sequence 1, Application US/08457822 | 2.22e-02 |
| 19 | 52 | 45.6 | 367 | US-08-948- Sequence 4, Application US/08457822 | 2.22e-02 |
| 20 | 52 | 45.6 | 367 | US-08-075- Sequence 2, Application US/08457822 | 2.22e-02 |
| 21 | 52 | 45.6 | 367 | US-08-075- Sequence 2, Application US/08457822 | 2.22e-02 |
| 22 | 52 | 45.6 | 367 | US-08-948- Sequence 2, Application US/08457822 | 2.22e-02 |
| 23 | 52 | 45.6 | 367 | US-08-948- Sequence 2, Application US/08457822 | 2.22e-02 |

| | | | | | | |
|----|----|------|------|---|---|----------|
| 24 | 52 | 45.6 | 375 | 1 | US-08-464- Sequence 22, Application US/08457822 | 2.22e-02 |
| 25 | 52 | 45.6 | 707 | 4 | PCI-US95-1 Sequence 2, Application US/08457822 | 2.22e-02 |
| 26 | 52 | 45.6 | 707 | 2 | US-08-663- Sequence 2, Application US/08457822 | 2.22e-02 |
| 27 | 52 | 45.6 | 797 | 2 | US-08-323- Sequence 1, Application US/08457822 | 2.22e-02 |
| 28 | 52 | 45.6 | 1012 | 3 | US-09-031- Sequence 4, Application US/08457822 | 2.22e-02 |
| 29 | 52 | 45.6 | 1012 | 3 | US-09-031- Sequence 4, Application US/08457822 | 2.22e-02 |
| 30 | 52 | 45.6 | 1012 | 3 | US-08-219- Sequence 19, Application US/08457822 | 2.22e-02 |
| 31 | 52 | 45.6 | 1012 | 3 | US-08-219- Sequence 3, Application US/08457822 | 2.22e-02 |
| 32 | 52 | 45.6 | 1012 | 3 | US-09-031- Sequence 3, Application US/08457822 | 2.22e-02 |
| 33 | 52 | 45.6 | 1012 | 3 | US-09-031- Sequence 3, Application US/08457822 | 2.22e-02 |
| 34 | 52 | 45.6 | 1012 | 3 | US-09-031- Sequence 3, Application US/08457822 | 2.22e-02 |
| 35 | 52 | 45.6 | 1012 | 3 | US-09-031- Sequence 3, Application US/08457822 | 2.22e-02 |
| 36 | 52 | 45.6 | 1012 | 3 | US-09-031- Sequence 3, Application US/08457822 | 2.22e-02 |
| 37 | 52 | 45.6 | 1012 | 3 | US-09-031- Sequence 3, Application US/08457822 | 2.22e-02 |
| 38 | 52 | 45.6 | 1012 | 3 | US-09-031- Sequence 3, Application US/08457822 | 2.22e-02 |
| 39 | 52 | 45.6 | 1012 | 3 | US-09-031- Sequence 3, Application US/08457822 | 2.22e-02 |
| 40 | 52 | 45.6 | 1012 | 3 | US-09-031- Sequence 3, Application US/08457822 | 2.22e-02 |
| 41 | 52 | 45.6 | 1012 | 3 | US-09-031- Sequence 3, Application US/08457822 | 2.22e-02 |
| 42 | 52 | 45.6 | 1012 | 3 | US-09-031- Sequence 3, Application US/08457822 | 2.22e-02 |
| 43 | 52 | 45.6 | 1012 | 3 | US-09-031- Sequence 3, Application US/08457822 | 2.22e-02 |
| 44 | 52 | 45.6 | 1012 | 3 | US-09-031- Sequence 3, Application US/08457822 | 2.22e-02 |
| 45 | 52 | 45.6 | 1895 | 2 | US-08-619- Sequence 4, Application US/08457822 | 2.22e-02 |

ALIGNMENTS

RESULT 1

ID US-08-467-822-16 STANDARD: PRI: 93 AA.

AC xxxxxx

DE

XX Sequence 36, Application US/08457822

XX Sequence 36, Application US/08457822

CC Patent No. 5843460

CC GENERAL INFORMATION:

CC APPLICANT: Sabigne, Andr

CC APPLICANT: Sauerbaum, S

CC APPLICANT: Thibierge, Jean-Miche

CC TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST

CC TITLE OF INVENTION: HELICOBACTER INFECTION POLYPEPTIDES FOR USE IN THE

CC TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID

CC NUMBER OF SEQUENCES: 44

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Pharmacia, 1100 I Street, N.W.

CC ALPHABET: 1370 I Street, N.W.

CC CITY: Washington

CC STATE: D.C.

CC COUNTRY: USA

CC ZIP: 20005-3315

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Floppy disk

CC COMPUTER: IBM PC compatible

CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: Patent In Release #1.0, Version #1.50

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/467,822

CC FILING DATE: 06-JUN-1995

CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/447,177

CC FILING DATE: 19-MAY-1995

CC CLASSIFICATION: 435

CC PRIOR APPLICATION DATA:

CC APPLICATION NUMBER: US 08/432,697

CC FILING DATE: 02-MAY-1995

CC CLASSIFICATION: 435

CC ATTORNEY/AGENT INFORMATION:


```

CC TYPE: AMINO ACID
CC STRANDEDNESS: UNKNOWN
CC TOPOLOGY: UNKNOWN
SQ SEQUENCE 612 AA: 68803 MW: 1972484 CN:

Query Match 48.2% Score 55: DB 2: Length 612:
Best Local Similarity 33.3% Pred. No. 1: 19e-02:
Matches 5: Conservative 7: Mismatches 3: Indels 0: Gaps 0:

DE 182 FLVIAAIAIAIAIFX 196
CY 3 YLLSARDVLAVVSK 17

RESULT 11
DE US-08-162-809-8 STANDARD: PRI: 973 AA.
XX
XX xxxxxx
DE
DE Sequence 8, Application US/08/62809
XX Sequence 8, Application US/08/62809
XX Patent No. 5457048
XX GENERAL INFORMATION:
XX APPLICANT: Pasquale, Elena B.
XX TITLE OF INVENTION: NOVEL EPH-RELATED TYROSINE KINASES,
XX TITLE OF INVENTION: NUCLEOTIDE SEQUENCES, AND METHODS OF USE
XX NUMBER OF SEQUENCES: 26
XX CORRESPONDENCE ADDRESS:
XX ADDRESSEE: CAMPBELL AND FLORES
XX STREET: 4470 La Jolla Village Drive, Suite 700
XX CITY: San Diego
XX STATE: California
XX COUNTRY: United States of America
XX ZIP: 92122
XX COMPUTER READABLE FORM:
XX MEDIUM TYPE: Floppy disk
XX COMPUTER: IBM PC compatible
XX OPERATING SYSTEM: PC-DOS/MS-DOS
XX SOFTWARE: Patent in Release #1.25
XX CURRENT APPLICATION DATA:
XX APPLICATION NUMBER: US/08/162,809
XX FILING DATE:
XX CLASSIFICATION:
XX ATTORNEY/AGENT INFORMATION:
XX NAME: Campbell, Graham A.
XX REGISTRATION NUMBER: 3,1915
XX REFERENCE/DOCKET NUMBER: P-03 9503
XX TELECOMMUNICATION INFORMATION:
XX TELEPHONE: (619) 535-9001
XX TELEFAX: (619) 535-8945
XX INFORMATION FOR SEQ ID NO: 8:
XX SEQUENCE CHARACTERISTICS:
XX LENGTH: 973 amino acids
XX TYPE: amino acid
XX TOPOLOGY: linear
XX MOLECULE TYPE: protein
XX SEQUENCE 973 AA: 108597 MW: 4992047 CN:

Query Match 48.2% Score 55: DB 1: Length 973:
Best Local Similarity 33.3% Pred. No. 1: 19e-02:
Matches 5: Conservative 7: Mismatches 3: Indels 0: Gaps 0:

DE 543 FLVIAAIAIAIAIFX 557
CY 3 YLLSARDVLAVVSK 17

RESULT 12
DE PCT-US-95-10194-35 STANDARD: PRI: 327 AA.
XX
XX xxxxxx
DE
DE Sequence 35, Application US/0842235B
XX Sequence 35, Application US/0842235B
XX Patent No. 5801042
XX GENERAL INFORMATION:
XX APPLICANT: Chang, Yuan
XX APPLICANT: Moore, Patrick S.
XX TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
XX TITLE OF INVENTION: SEQUENCES AND USES THEREOF
XX NUMBER OF SEQUENCES: 47
XX CORRESPONDENCE ADDRESS:
XX ADDRESSEE: Cooper & Dunham LLP

```

CC STREET: 1185 Avenue of the Americas
CC City: New York
CC STATE: New York
CC COUNTRY: U.S.A.
CC ZIP: 10036
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent in Release #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/420.235B
CC FILING DATE:
CC CLASSIFICATION: 424
CC ATTORNEY/AGENT INFORMATION:
CC NAME: White, Long P.
CC REGISTRATION NUMBER: 28,679
CC REFERENCE/DOCKET NUMBER: 45185-B
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 278-0400
CC TELEFAX: (212) 391-0525
CC INFORMATION FOR SEQ ID NO: 35:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 327 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC CLONE: Consensus
CC SEQUENCE 327 AA: 36122 MW: 55627. GN:
CC
CC Query Match 47.4% Score 54: DB 1: Length 327:
CC Best Local Similarity 31.3% Pred. No. 1.47e-02:
CC Matches 5: Conservative 9: Mismatches 2: Indels 0: Gaps 0:
CC
CC 303 DYLIIPSYDIPAITM 318
CC 311 111 111 111
CC 2 EYLSARDVLAVSK 17
CC
CC
CC RESULT 14
CC US-08-745 934-1 STANDARD: PRT: 572 AA.
CC XXXXXX
CC
CC Sequence 1, Application US/08/745934
CC Patent No. 5861466
CC GENERAL INFORMATION:
CC APPLICANT: Hillman, Jennifer L.
CC APPLICANT: Hawkins, Philip P.
CC TITLE OF INVENTION: HUMAN SQUALENE EPOXIDASE
CC NUMBER OF SEQUENCES: 4
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Incyte Pharmaceuticals, Inc.
CC STREET: 3174 Porter Drive
CC City: Palo Alto
CC STATE: California
CC COUNTRY: USA
CC ZIP: 94304
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette
CC COMPUTER: IBM compatible
CC OPERATING SYSTEM: DOS
CC SOFTWARE: FASTSEQ Version 1.5
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/745.934
CC FILING DATE: Herewith
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:

CC NAME: Billings, Lucy J.
CC REGISTRATION NUMBER: 36,749
CC REFERENCE/DOCKET NUMBER: PF-015: US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (415) 855-0555
CC TELEFAX: (415) 845-4166
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 572 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC IMMEDIATE SOURCE:
CC LIBRARY: Consensus
CC CLONE: Consensus
CC SEQUENCE 572 AA: 63940 MW: 177482. GN:
CC
CC Query Match 47.4% Score 54: DB 2: Length 572:
CC Best Local Similarity 35.7% Pred. No. 1.47e-02:
CC Matches 5: Conservative 5: Mismatches 4: Indels 3: Gaps 0:
CC
CC DB 19 DFITLAKREVLLCV 32
CC 111 111 111
CC 2 EYLSARDVLAVV 15
CC
CC
CC RESULT 15
CC US-08-387-942C-4 STANDARD: PRT: 997 AA.
CC XXXXXX
CC
CC Sequence 4, Application US/08387942C
CC Patent No. 5939289
CC GENERAL INFORMATION:
CC APPLICANT: ERTESVAG, HELGA
CC APPLICANT: VALIA, SVEIN
CC APPLICANT: SKOK-BAKK, GUMUND
CC APPLICANT: LARSEN, BJORN
CC TITLE OF INVENTION: DNA COMPOUNDS COMPRISING SEQUENCES
CC TITLE OF INVENTION: ENCODING MANNURONAN C-5-EPIMERASE
CC NUMBER OF SEQUENCES: 52
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH, LLP
CC STREET: P.O. BOX 247
CC City: Falls Church
CC STATE: VA
CC COUNTRY: USA
CC ZIP: 22042
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent in Release #1.0. Version #1.1C
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/387.942C
CC FILING DATE: 09-MAY-1995
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: MURPHY JR, GERALD M.
CC REGISTRATION NUMBER: 28,977
CC REFERENCE/DOCKET NUMBER: 1809-106P
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 703-205-8000
CC TELEFAX: 703-205-8050
CC INFORMATION FOR SEQ ID NO: 4:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 997 amino acids
CC TYPE: amino acid

[M] [E] [S] [E] [R] [E] [D]
[] [] [] [] [] [] []
***** (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit,
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Musich_Pf protein - protein database search, using Smith-Waterman algorithm

Run On: Mon Jun 19 15:25:24 2000; MasPar time 15.35 Seconds

Tutorial output not generated.

Title: >US-09-142-524A-8

Description: (1-17) from US09142524A.pcp

Perfect Score: 114

Sequence: 1 EEWLILSARDVLAVVSK 17

Scoring table: PAM 150
Gap 15

Searched: 721208 seqs, 103765575 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database:
3-pending
18997 2:06 3:060 4:07 5:080 6:081 7:092 8:083 9:084A
10:084B 11:085 12:086 13:087 14:088 15:089 16:090 17:091
18:092 19:093 20:094 21:095 22:NEWP 23:NEWU60 24:NEWU8
25:NEWU9

Statistics: Mean 22.684; Variance 64.304; score 0.353

Pred No. is the number of results predicted by chance to have a
Score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

| SUMMARIES | | | | | | | |
|------------|-------|-------|--------|----|------------------|-------------------------------------|--|
| Result No. | Score | Match | Length | ID | Description | Pred. No. | |
| 1 | 114 | 100.0 | 93 | 10 | US-08-466-248-36 | Sequence 36, Application US/ 456248 | |
| 2 | 114 | 100.0 | 93 | 9 | US-08-452-107 | Sequence 36, Application 1,00e-04 | |
| 3 | 72 | 63.2 | 119 | 17 | US-09-107-137 | Sequence 36, Application 1,00e-01 | |
| 4 | 70 | 61.4 | 94 | 9 | US-08-412-134 | Sequence 38, Application 2,07e-01 | |
| 5 | 70 | 61.4 | 94 | 10 | US-08-456-134 | Sequence 38, Application 2,07e-01 | |
| 6 | 70 | 61.4 | 104 | 17 | US-09-134-134 | Sequence 4557, Application 2,07e-01 | |
| 7 | 66 | 57.9 | 111 | 20 | US-03-450-450 | Sequence 6153, Application 5,73e-01 | |
| 8 | 65 | 57.0 | 94 | 10 | US-08-466-134 | Sequence 39, Application 7,37e-01 | |
| 9 | 65 | 57.0 | 94 | 9 | US-08-432-134 | Sequence 39, Application 7,37e-01 | |
| 10 | 65 | 57.0 | 97 | 9 | US-08-432-134 | Sequence 40, Application 7,37e-01 | |
| 11 | 65 | 57.0 | 97 | 20 | US-09-472-134 | Sequence 40, Application 7,37e-01 | |
| 12 | 65 | 57.0 | 97 | 10 | US-08-466-134 | Sequence 40, Application 7,37e-01 | |
| 13 | 65 | 57.0 | 129 | 25 | US-05-252-134 | Sequence 178, Application 7,37e-01 | |
| 14 | 65 | 57.0 | 135 | 1 | PCT-US98-0 | Sequence 178, Application 7,37e-01 | |
| 15 | 65 | 57.0 | 135 | 14 | US-08-833-134 | Sequence 178, Application 7,37e-01 | |
| 16 | 63 | 55.3 | 388 | 18 | US-08-232-134 | Sequence 30, Application 1,21e-02 | |
| 17 | 61 | 53.5 | 118 | 9 | US-08-432-134 | Sequence 29, Application 1,99e-02 | |
| 18 | 61 | 53.5 | 118 | 10 | US-08-466-134 | Sequence 29, Application 1,99e-02 | |
| 19 | 61 | 53.5 | 154 | 17 | US-09-107-134 | Sequence 6000, Application 1,99e-02 | |
| 20 | 61 | 53.5 | 189 | 23 | US-60-198-134 | Sequence 1255, Application 1,99e-02 | |

| ALIGNMENTS | | | | | | | |
|------------|---------------------------|--|----------|-----|-------|--|--|
| RESULT | ID | US-08-466-248-36 | STANDARD | PRG | 93 AA | | |
| XX | AC | xxxxxx | | | | | |
| XX | CT | | | | | | |
| XX | DE | Sequence 36, Application US/ 456248 | | | | | |
| XX | CC | Sequence 36, Application US/ 456248 | | | | | |
| CC | GENERAL INFORMATION: | | | | | | |
| CC | APPLICANT: | Labigne, Andre | | | | | |
| CC | APPLICANT: | Sauerbaum, S | | | | | |
| CC | APPLICANT: | Ferrero, R | | | | | |
| CC | APPLICANT: | Truberge, J | | | | | |
| CC | TITLE OF INVENTION: | IMMUNICOMPOSITIONS AGAINST | | | | | |
| CC | TITLE OF INVENTION: | HEPATITIS INFECTION, POLYPEPTIDES FOR USE IN THE | | | | | |
| CC | TITLE OF INVENTION: | COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID | | | | | |
| CC | TITLE OF INVENTION: | POLYPEPTIDES | | | | | |
| CC | NUMBER OF SEQUENCES: | 44 | | | | | |
| CC | CORRESPONDENCE ADDRESS: | | | | | | |
| CC | ADDRESSEE: | Fitchman, D | | | | | |
| CC | ADDRESSEE: | Donner, N | | | | | |
| CC | STREET: | 1100 1 Street, N.W. | | | | | |
| CC | CITY: | Washington | | | | | |
| CC | STATE: | D.C. | | | | | |
| CC | COUNTRY: | USA | | | | | |
| CC | Zip: | 20005-3315 | | | | | |
| CC | COMPUTER READABLE FORM | | | | | | |
| CC | MEDIUM TYPE: | Floppy disk | | | | | |
| CC | COMPUTER: | IBM PC compatible | | | | | |
| CC | OPERATING SYSTEM: | PC-DOS/MS-DOS | | | | | |
| CC | SOFTWARE: | Patent Release #1.0, Version #1.30 | | | | | |
| CC | CURRENT APPLICATION DATA: | | | | | | |
| CC | APPLICATION NUMBER: | US/08/466,248 | | | | | |
| CC | FILING DATE: | 06-JUN-1995 | | | | | |
| CC | CLASSIFICATION: | 435 | | | | | |
| CC | PRIOR APPLICATION DATA: | | | | | | |
| CC | APPLICATION NUMBER: | US 08/447,177 | | | | | |
| CC | FILING DATE: | 19-MAY-1995 | | | | | |
| CC | CLASSIFICATION: | 435 | | | | | |
| CC | PRIOR APPLICATION DATA: | | | | | | |
| CC | APPLICATION NUMBER: | US 08/432,697 | | | | | |
| CC | FILING DATE: | 02-MAY-1995 | | | | | |

CC CLASSIFICATION: 435
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Meyers, Kenneth J.
 CC REGISTRATION NUMBER: 25,146
 CC REFERENCE/DOCKET NUMBER: 03495, 0137-00000
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (202) 408-4000
 CC TELEFAX: (202) 408-4400
 CC INFORMATION FOR SEQ ID NO: 36:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 93 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS: single
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC SEQUENCE: 93 AA: 10002 MW: 48030 GN.
 CC
 CC Query Match 100.0% Score 114: DB 12: Length 93:
 CC Best Local Similarity 100.0% Pident. No. 1 00e-04:
 CC Matches 17: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 CC
 CC 77 EYLLSARDVLAVSK 93
 CC 1 EYLLSARDVLAVSK 17
 CC
 CC RESULT 2
 CC US-09-142-524A-8.rap STANDARD: PFI: 93 AA.
 CC
 CC xxxxxx
 CC
 CC Sequence 36, Application 05/08/12697
 CC
 CC Sequence 36, Application: US/08432697
 CC GENERAL INFORMATION:
 CC APPLICANT: Labigne, Anes
 CC APPLICANT: Sauerbaum, Sebastian
 CC APPLICANT: Ferrero, Richard L.
 CC APPLICANT: Thibierge, Jean-Michel
 CC TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
 CC TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
 CC TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
 CC TITLE OF INVENTION: POLYPEPTIDES
 CC NUMBER OF SEQUENCES: 44
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Finckh, Henderson, Parabow, Garrett &
 CC ADDRESSEE: Dundee
 CC STREET: 1400 15th St., N.W.
 CC CITY: Washington
 CC STATE: D.C.
 CC COUNTRY: USA
 CC ZIP: 20005-3315
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC OPERATING SYSTEM: IBM PC compatible
 CC SOFTWARE: Patent Release #1.0, Version #1.30
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/432,697
 CC FILING DATE: 02-MAY-1995
 CC CLASSIFICATION: 424
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Meyers, Kenneth J.
 CC REGISTRATION NUMBER: 25,146
 CC REFERENCE/DOCKET NUMBER: 03495, 0137-00000
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (202) 408-4000
 CC TELEFAX: (202) 408-4400
 CC INFORMATION FOR SEQ ID NO: 36:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 93 amino acids

CC TYPE: amino acid
 CC STRANDEDNESS: single
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC SEQUENCE: 93 AA: 10002 MW: 48030 GN.
 CC
 CC Query Match 100.0% Score 114: DB 9: Length 93:
 CC Best Local Similarity 100.0% Pident. No. 1 00e-04:
 CC Matches 17: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 CC
 CC 77 EYLLSARDVLAVSK 93
 CC 1 EYLLSARDVLAVSK 17
 CC
 CC RESULT 3
 CC US-09-107-532-6969 STANDARD: PFI: 11 AA
 CC
 CC xxxxxx
 CC
 CC Sequence 6969, Application 05/0917512
 CC
 CC Sequence 6969, Application: US/09107512
 CC GENERAL INFORMATION:
 CC APPLICANT: Lynn A. Co. Collette-Stamm and David Bush
 CC TITLE OF INVENTION: NOVEL ACID AND AMINO ACID SEQUENCES RELATING TO
 CC TITLE OF INVENTION: ENTEROCOCCUS FAECIUM FOR
 CC TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
 CC NUMBER OF SEQUENCES: 7408
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: GENOME THERAPEUTICS CORPORATION
 CC STREET: 100 Beaver Street
 CC CITY: Waltham
 CC STATE: Massachusetts
 CC COUNTRY: USA
 CC ZIP: 02154
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: CD-ROM 1504560
 CC OPERATING SYSTEM:
 CC SOFTWARE:
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/ 4/107,532
 CC FILING DATE:
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 1 125598
 CC FILING DATE: May 14 1997
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 1 504560
 CC FILING DATE: July 2 1997
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Attelleo, Pamela Debra
 CC REGISTRATION NUMBER: 40,489
 CC REFERENCE/DOCKET NUMBER: GTC-012
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (781)893-5907
 CC TELEFAX: (781)893-8277
 CC INFORMATION FOR SEQ ID NO: 6969:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 119 amino acids
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC HYPOTHETICAL: YES
 CC ORIGINAL SOURCE:
 CC ORGANISM: Enterococcus faecium
 CC FEATURE:
 CC NAME/KEY: misc_feature
 CC LOCATION: 1...119
 CC SEQUENCE: 119 AA: 12807 MW: 74418 GN


```

DE XX Sequence 4557, Application US/09134000A
DE XX
DE XX Sequence 4557, Application US/09134000A
DE XX
DE XX GENERAL INFORMATION:
DE XX APPLICANT: Lynn Doucette-Stamm et al
DE XX TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS
DE XX TITLE OF INVENTION: FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS
DE XX FILE REFERENCE: CTC-005
DE XX CURRENT APPLICATION NUMBER: US/09/134.000A
DE XX CURRENT FILING DATE: 1998-08-13
DE XX NUMBER OF SEQ ID NOS: 6810
DE XX SEQ ID NO 4557
DE XX LENGTH: 104
DE XX TYPE: PRT
DE XX ORGANISM: Enterococcus faecalis
DE XX SEQUENCE 104 AA: 11303 MW: 63709 CN:

Query Match 51.4% Score 70: DB 17: Length 104:
Best Local Similarity 54.3% Pred. No. 2.07e-01:
Matches 9: Conservative 4: Mismatches 1: Indels 0: Gaps 0:

DE 90 EYLTIVAKDILATV 103
QY 2 EYLTIVAKDILATV 15

RESULT 7
DE US-09-450-969-616 STANDARD: PRT: 111 AA.
DE XX
DE XX
DE XX
DE XX
DE XX Sequence 6163, Application US/09450969
DE XX
DE XX Sequence 6163, Application US/09450969
DE XX
DE XX GENERAL INFORMATION:
DE XX APPLICANT: Lynn Doucette-Stamm et al
DE XX TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
DE XX TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
DE XX FILE REFERENCE: PATH99-09A
DE XX CURRENT APPLICATION NUMBER: US/09/450.969
DE XX CURRENT FILING DATE: 1999-11-29
DE XX NUMBER OF SEQ ID NOS: 7544
DE XX SEQ ID NO 6163
DE XX LENGTH: 111
DE XX TYPE: PRT
DE XX ORGANISM: Staphylococcus epidermidis
DE XX SEQUENCE 111 AA: 12144 MW: 11750 IN

Query Match 57.9% Score 60: DB 20: Length 111:
Best Local Similarity 53.8% Pred. No. 5.73e-01:
Matches 7: Conservative 4: Mismatches 2: Indels 0: Gaps 0:

DE 98 YLILNEEDILATV 110
QY 3 YLILNEEDILATV 15

RESULT 8
DE US-08-455-243-39 STANDARD: PRT: 94 AA.
DE XX
DE XX
DE XX
DE XX
DE XX Sequence 39, Application US/08456248
DE XX
DE XX Sequence 39, Application US/08456248
DE XX
DE XX GENERAL INFORMATION:
DE XX APPLICANT: Labigne, Agnes
DE XX APPLICANT: Sauerbaum, Sebastian
DE XX APPLICANT: Ferrero, Richard L.
DE XX APPLICANT: Thiberge, Jean-Michel

```

```

CC APPLICANT: Sauerbaum, Sebastian
CC APPLICANT: Ferrero, Richard L.
CC APPLICANT: Thiberge, Jean-Michel
CC TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
CC TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
CC TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
CC TITLE OF INVENTION: POLYPEPTIDES
CC NUMBER OF SEQUENCES: 44
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
CC STREET: 3300 I Street, N.W.
CC CITY: Washington
CC STATE: D.C.
CC COUNTRY: USA
CC ZIP: 20005-3315
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release 4.1.0, Version 4.1.0
CC CURRENT APPLICATION DATA: US/09/466.248
CC APPLICATION NUMBER: US/09/466.248
CC FILING DATE: 06-JUN-1995
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/447,177
CC FILING DATE: 19-MAY-1995
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/432,697
CC FILING DATE: 02-MAY-1995
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Meyers, Kenneth J.
CC REGISTRATION NUMBER: 25,146
CC REFERENCE/DOCKET NUMBER: 03495 0137-02000
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (202) 408-4000
CC TELEFAX: (202) 408-4400
CC INFORMATION FOR SEQ ID NO: 19:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 94 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: Single
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 94 AA: 10308 MW: 46636 CN:

Query Match 57.0% Score 60: DB 19: Length 94:
Best Local Similarity 53.8% Pred. No. 5.73e-01:
Matches 9: Conservative 4: Mismatches 4: Indels 0: Gaps 0:

DE 79 EYLTIVAKDILATV 93
QY 1 EYLTIVAKDILATV 15

RESULT 9
DE US-06-432-637-39 STANDARD: PRT: 94 AA.
DE XX
DE XX
DE XX
DE XX
DE XX Sequence 39, Application US/06-432697
DE XX
DE XX Sequence 39, Application US/06-432697
DE XX
DE XX GENERAL INFORMATION:
DE XX APPLICANT: Labigne, Agnes
DE XX APPLICANT: Sauerbaum, Sebastian
DE XX APPLICANT: Ferrero, Richard L.
DE XX APPLICANT: Thiberge, Jean-Michel

```

```

CC      TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
CC      TITLE OF INVENTION: HELICOBACTER INFECTION. POLYPEPTIDES FOR USE IN THE
CC      TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
CC      TITLE OF INVENTION: POLYPEPTIDES
CC      NUMBER OF SEQUENCES: 44
CC      CORRESPONDENCE ADDRESS:
CC      ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
CC      ADDRESSEE: Dunner
CC      STREET: 1300 I Street, N.W.
CC      CITY: Washington
CC      STATE: D.C.
CC      COUNTRY: USA
CC      ZIP: 20005-3315
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: PatentIn Release #1.0, Version #1.30
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: US/08/432,597
CC      FILING DATE: 02-MAY-1995
CC      CLASSIFICATION: 424
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: Meyers, Kenneth J.
CC      REGISTRATION NUMBER: 25,146
CC      REFERENCE/DOCKET NUMBER: 03495-0137-00000
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: (202) 408-4000
CC      TELEFAX: (202) 408-4400
CC      INFORMATION FOR SEQ ID NO: 39:
CC      SEQUENCE CHARACTERISTICS:
CC      LENGTH: 94 amino acids
CC      TYPE: amio acid
CC      STRANDEDNESS: single
CC      TOPOLOGY: linear
CC      MOLECULE TYPE: protein
CC      SEQUENCE 94 AA: 10308 MW: 46636 CN:
CC
CC      Query Match 57.0% Score 65 DB 9: Length 94:
CC      best Local Similarity 60.0% Pred. No. 7.37e+01:
CC      Matches 9: Conservative 2: Mismatches 4: Indels 0: Gaps 0:
CC
CC      DB 75 EHYLLKRDQGLAVV 93
CC      Q7 2 EHYLLSARDVLAVV 15
CC
CC      RESUME 10
CC      SEQ US-08-432-597-40 STANDARD: PRT: 97 AA.
CC      XX xxxxxx
CC
CC      Sequence 40, Application: US/08432697
CC
CC      Sequence 40, Application US/08432697
CC      GENERAL INFORMATION:
CC      APPLICANT: Labigne, Agnes
CC      APPLICANT: Sauerbaum, Sebastian
CC      APPLICANT: Ferrero, Richard L.
CC      APPLICANT: Thiberge, Jean-Michel
CC      TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
CC      TITLE OF INVENTION: HELICOBACTER INFECTION. POLYPEPTIDES FOR USE IN THE
CC      TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
CC      TITLE OF INVENTION: POLYPEPTIDES
CC      NUMBER OF SEQUENCES: 44
CC      CORRESPONDENCE ADDRESS:
CC      ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
CC      ADDRESSEE: Dunner
CC      STREET: 1300 I Street, N.W.
CC      CITY: Washington
CC      STATE: D.C.

```


CC ATTORNEY/AGENT INFORMATION:
CC NAME: Clark, Paul T.
CC REGISTRATION NUMBER: 30,162
CC REFERENCE/DOCKET NUMBER: 06132/041WO1
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 617-428-0200
CC TELEFAX: 617-428-7045
CC TELEX:

CC INFORMATION FOR SEQ ID NO: 178:

CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 136 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear

CC MOLECULE TYPE: protein
CC FRAGMENT TYPE: internal
CC FRAGMENT TYPE: internal

CC SEQUENCE 136 AA: 15614 MW: 100717 CN:

CC Query Match 57.0% Score 65; DB 1; Length 136;
CC Best Local Similarity 50.3%; Pred. NO. 7.37e-61;
CC Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

CC 8 EDYMLMSREVSFAVG 23
CC 136 136 136 136 136
CC 1 EBYLLSARDVLAWS 16

CC RESULT 15
CC US-08-633-457-179 STANDARD: PRI: 136 AA.
CC XXXXXX

CC Sequence 178. Application US/09833457

CC Sequence 179. Application US/08833457
CC GENERAL INFORMATION:

CC APPLICANT: Harold Kleanthous et al.

CC TITLE OF INVENTION: Identification of Polynucleotides

CC TITLE OF INVENTION: Encoding Novel Helicobacter Polypeptides in the Helicobacter

CC TITLE OF INVENTION: Genome

CC NUMBER OF SEQUENCES: 370

CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Clark & Biring LLP

CC STREET: 176 Federal Street

CC CITY: Boston

CC STATE: MA

CC COUNTRY: USA

CC ZIP: 02110

CC COMPUTER READABLE FORM:

CC MEDIUM TYPE: Diskette

CC COMPUTER: IBM Compatible

CC OPERATING SYSTEM: DOS

CC SOFTWARE: FASTSEQ for Windows Version 2.0

CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/08/833,457

CC FILING DATE: 01-APR-97

CC CLASSIFICATION: 435

CC ATTORNEY/AGENT INFORMATION:

CC NAME: Clark, Paul T.

CC REGISTRATION NUMBER: 30,162

CC REFERENCE/DOCKET NUMBER: 06132/041001

CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: 617-428-0200

CC TELEFAX: 617-428-7045

CC TELEX:

CC INFORMATION FOR SEQ ID NO: 178:

CC SEQUENCE CHARACTERISTICS:

CC LENGTH: 136 amino acids

CC TYPE: amino acid

CC STRANDEDNESS: single

CC TOPOLOGY: linear

CC MOLECULE TYPE: protein
CC FRAGMENT TYPE: internal
CC SEQUENCE 136 AA: 15614 MW: 100717 CN:

CC Query Match 57.0% Score 65; DB 1; Length 136;
CC Best Local Similarity 50.3%; Pred. NO. 7.37e-61;
CC Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

CC 8 EDYMLMSREVSFAVG 23
CC 136 136 136 136 136
CC 1 EBYLLSARDVLAWS 16

CC Search completed: Mon Jun 19 16:12:45 2000
CC Job time : 21 secs.

M A P E N E

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit,
Copyright (c) 1994-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd
Msrch_pp protein - protein database search using Smith-Waterman algorithm
Run on: Mon Jun 19 16:08:02 2000: Mispair time 5.39 Seconds
Tabular output not generated. 148,858 Million cell updates/sec

Title: >US-09-142-524A-8
Description: (1-17) from US09.42524A.pep
Perfect Score: 114
Sequence: 1 EEWLLSARDVLAVVSK 17
Scoring table: PAM 150
Gap 15
Searched: 142080 seqs, 47172406 residues
Post-Processing: Minimum Match 0%
Listing first 45 summaries
Database: pir62
1:pir1 2:pir2 3:pir3 4:pir4
Statistics: Mean 29.138; Variance 41.997; scale 0.694
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

| SUMMARIES | | | | | | | |
|-----------|-----|-------|--------------|----|--------|-----------------------|-----------|
| Result | No. | Score | Match Length | SN | ID | Description | Pred. No. |
| 1 | 14 | 100.0 | 100 | 2 | S25180 | Heat shock protein cr | 1.07e-10 |
| 2 | 114 | 100.0 | 100 | 2 | CH2556 | 10K T-cell antigen - | 1.07e-10 |
| 3 | 114 | 100.0 | 100 | 2 | BWY8A | Chaperonin groES - My | 1.07e-10 |
| 4 | 114 | 100.0 | 169 | 2 | S72818 | Heat shock protein ch | 1.07e-10 |
| 5 | 105 | 92.1 | 102 | 2 | S37565 | Chaperonin groES - St | 1.16e-08 |
| 6 | 103 | 90.4 | 102 | 2 | A41325 | Heat shock protein 18 | 3.24e-08 |
| 7 | 80 | 70.2 | 103 | 2 | A35721 | groES protein - Synec | 2.39e-03 |
| 8 | 80 | 70.2 | 103 | 1 | BVFCGS | Chaperonin groES - Sy | 2.39e-03 |
| 9 | 79 | 69.3 | 100 | 1 | BWY7B | Chaperonin groES - My | 3.79e-03 |
| 10 | 78 | 58.4 | 103 | 2 | T0683C | Chaperonin groES - Cy | 5.97e-03 |
| 11 | 76 | 66.7 | 96 | 2 | S34937 | Heat shock protein hs | 1.47e-02 |
| 12 | 75 | 65.8 | 94 | 2 | A49855 | Heat shock protein Gr | 2.30e-02 |
| 13 | 75 | 65.8 | 94 | 2 | A41884 | Heat shock protein (C | 2.30e-02 |
| 14 | 74 | 64.9 | 94 | 2 | JG1479 | Heat shock protein TG | 3.58e-02 |
| 15 | 73 | 64.0 | 88 | 2 | PC4238 | Heat shock protein Gr | 9.56e-02 |
| 16 | 73 | 64.0 | 94 | 2 | S72613 | Chaperonin groES - Th | 5.56e-02 |
| 17 | 71 | 62.3 | 94 | 2 | S32105 | Chaperonin groES - La | 1.33e-01 |
| 18 | 71 | 62.3 | 94 | 2 | JN0660 | Heat shock protein gr | 1.33e-01 |
| 19 | 70 | 61.4 | 94 | 2 | CN0500 | Heat shock protein 10 | 2.04e-01 |
| 20 | 70 | 61.4 | 96 | 2 | A49203 | Heat shock protein Gr | 2.04e-01 |
| 21 | 69 | 60.5 | 106 | 2 | S77563 | Chaperonin groES - Sy | 3.12e-01 |
| 22 | 58 | 59.6 | 96 | 1 | A42281 | Symbiotic sys - pea | 4.77e-01 |
| 23 | 67 | 58.8 | 94 | 2 | S68248 | Chaperonin groES - Cl | 7.24e-01 |

ALIGNMENTS

```
RESULT 1
ENTRY 1 S25180 #type complete
TITLE heat shock protein groES - Mycobacterium leprae
ALTERNATE_NAMES 10K chaperonin
ORGANISM Mycobacterium leprae
DATE 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change
ACCESSIONS S25180; S72897
REFERENCE de Wit, T.F.R.; Beketic, S.; Osland, A.; Miko, T.; L. Hezans,
P.W.M.; van Soelingen, D.; Drijfhout, J.; Schoenmaker, R.;
Janson, A.A.; Thole, J.E.R.
Mol. Microbiol. 1992; 6:1995-2007
Mycobacteria contain two groES genes: the second
with groES.
#cross-references M101:92374850
#accession S25180
#molecule_type DNA
#residues 1-103 #label DEW
#cross-references EMBL:Z11665
#note the authors translated the initiation codon GTC to:
residue 1 as Val.
REFERENCE S74580
#authors Smith, D.K.; Robinson, K
#accession Submitted to the EMBL data library, November 1994
#description Mycobacterium leprae, strain H37Rv
#accession S72997
#status Preliminary
#molecule_type DNA
#residues 1-100 #label SM1
#cross-references EMBL:U00020; NID:q467102; PIDN:AAA173111;
PID:q467129
GENETICS
#gene groES: chpA
#start_codon GTC
#ASSIGNATION #superfamily chaperonin groES
KEYWORDS heat shock; stress-induced protein
SUMMARY #length 100 #molecular_weight 10800 #checksum 7362
Query Match 100.0% Score 114; DB 2; Length 100;
Best Local Similarity 100.0% Pred. No. 1.07e-10;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Do 84 EEWLLSARDVLAVVSK 100
|||||:|||||
Qy 1 EEWLLSARDVLAVVSK 17
```

```

RESULT      2
ENTRY
  TITLE      JH0556      *type complete
  ID         JH0556      *type complete
  ORGANSIM   JH0556      *type complete
  DATE       12-Sep-1997
  REFERENCE  JH0556
  AUTHORS    Mehra, V.; Bloom, B.R.; Bajardi, A.C.; Grisso, C.L.; Sieling,
              P.A.; Alland, D.; Conville, J.; Fan, X.; Hunter, S.W.;
              Brennan, P.J.; Rea, T.H.; Modlin, R.L.
  JOURNAL    J. Exp. Med. (1992) 175:275-284
  TITLE      A major T cell antigen of Mycobacterium leprae is a 10-kD
              heat-shock cognate protein.
  CROSS-REFERENCES EMBL:X89413
  SUMMARY    *length 100 *molecular-weight 10766 *checksum 7372
              *Query Match 100.0% Score 114; DB 2; Length 100;
              *Best Local Similarity 100.0%; Pred. No. 1,07e-10;
              *Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB      84 EBYLILSARDVLAVSK 100
Q      1 EBYLILSARDVLAVSK 17
      1 |||||.....|||
      2 BWMYRA      *type complete
      3 chaperonin groES - Mycobacterium tuberculosis
      4 10K antigen; BCG-a homolog; Cpn10; heat shock protein 10K
      5 (tspl0); immunogenic protein; BCG-a
      6 *formal-name Mycobacterium tuberculosis
      7 31-Dec-1990 *sequence-revision 31-Dec-1990 *text-change
      8 16-Jul-1999
      9 ACCESSIONS S03381; S07272; A47292; G70737; A46481; S18040
      10 REFERENCE S03381
      11 AUTHORS Baird, P.N.; Hall, L.M.C.; Coates, A.R.M.
      12 Nucleic Acids Res. (1988) 16:9347
      13 A major antigen from Mycobacterium tuberculosis which is
      14 homologous to the heat shock proteins groES from E. coli
      15 and the hsp90 gene product of Coxidia burnetii.
      16 *cross-references M01D:6401584
      17 *accession S01381
      18 *molecule_type DNA
      19 *residues 1-100 *label BA1
      20 *cross-references EMBL:X12598; NID:9444571; PID:5581358
      21 REFERENCE A37166
      22 AUTHORS Baird, P.N.; Hall, L.M.C.; Coates, A.R.M.
      23 J. Gen. Microbiol. (1989) 135:931-939
      24 *journal Cloning and sequence analysis of the 10 kDa antigen gene of
      25 Mycobacterium tuberculosis.
      26 *cross-references M01D:90095443
      27 *accession A37166
      28 *status preliminary
      29 *molecule_type DNA
      30 *residues 1-100 *label BA2
      31 *cross-references GB:M25258; GB:X12598
      32 REFERENCE S02727
      33 AUTHORS Shinnick, T.M.; Plakoytis, B.B.; Hyshe, A.D.; van Landingham,
              R.M.; Walker, L.L.
      34 Nucleic Acids Res. (1989) 17:1254
      35 *journal The Mycobacterium tuberculosis BCG-a protein has homology
      36 with the Escherichia coli GroES protein.
      37 *cross-references M01D:89160258
      38
      39 *accession S02727
      40 *molecule_type DNA
      41 *residues 1-100 *label SH1
      42 *cross-references EMBL:X13739; NID:9444571; PID:CAA12003.1; PID:5581360
      43 REFERENCE A47292
      44 AUTHORS Kong, I.H.; Coates, A.R.; Butcher, P.D.; Hickman, C.D.;
              Shinnick, T.M.
      45 *journal Proc. Natl. Acad. Sci. U.S.A. (1993) 90:2608-2612
      46 *title Mycobacterium tuberculosis expresses two chaperonin-like
              homologs.
      47 *cross-references M01D:93219382
      48 *accession A47292
      49 *status preliminary
      50 *molecule_type DNA
      51 *residues 1-100 *label X0N
      52 *cross-references EMBL:X60560; NID:9444599; PID:CAA42982.1; PID:5581363
      53 *note Sequence extracted from NDB: Backbone (NDB:1LA9955)
      54
      55 REFERENCE A70500
      56 *accession Cole, S.T.; Brosch, P.; Parkhill, J.; Garnier, J.; Churcher,
              C.; Harris, D.; Gordon, S.V.; Eigmeier, K.; Gas, S.; Barry,
              III, C.E.; Tekle, E.; Haddock, K.; Basham, D.; Brown, D.;
              Chillingworth, T.; Connor, K.; Davies, K.; Devlin, K.;
              Feltham, D.; Gentles, S.; Hamlin, N.; Holroyd, S.;
              Hornsby, L.; Jager, K.; Krogh, A.; McLean, C.; Mulle, S.;
              Murphy, L.; Oliver, S.; Osborne, J.; Quail, M.A.;
              Rajandream, M.J.; Rogers, J.; Rutter, S.; Seeger, K.;
              Skelton, S.; Squares, S.; Soares, R.; Sulten, J.;
              Taylor, K.; Whitehead, S.; Barrett, B.G.
      57 *journal Nature (1998) 393:537-544
      58 *title Deciphering the biology of Mycobacterium tuberculosis from
              the complete genome sequence.
      59 *cross-references M01D:98295987
      60 *accession G70737
      61 *status nucleic acid sequence not shown; translation not shown
      62 *molecule_type DNA
      63 *residues 1-100 *label G0U
      64 *cross-references GB:Z77265; GB:A123456; NID:93761609;
              PID:CAA60055.1; PID:62591809; PID:449369
      65 *exper mental-source strain H37Rv
      66 REFERENCE A46481
      67 AUTHORS Barcus, P.; Fehra, V.; Pivoire, B.; Fong, S.; Brennan,
              P.; Voegtline, M.S.; Minden, P.; Houghton, R.A.; Bloom,
              B.R.; Modlin, R.L.
      68 *journal J. Immunol. (1992) 148:1835-1840
      69 *title Immunoreactivity of a 10-kDa antigen of Mycobacterium
              tuberculosis.
      70 *cross-references M01D:92176646
      71 *accession A46481
      72 *status preliminary
      73 *molecule_type protein
      74 *residues 233 *label: BAW
      75 *note Sequence extracted from NDB: Backbone (NDB:1K7128)
      76
      77 GENETICS
      78 *gene groES; cpn10
      79 *start_codon GIG
      80 *complex functional chaperonin includes 14 chains of groEL and 7 of
              groES
      81
      82 FUNCTION
      83 *description mediates protein folding and renaturation
      84 *classification *superfamily chaperonin groES
      85 *keywords ATP; heat shock; molecular chaperone; stress-induced protein
      86 FEATURE
      87 2-100
      88
      89 SUMMARY
      90 *length 100 *molecular-weight 10804 *checksum 7634
      91
      92 Query Match 100.0%; Score 114; DB 1; Length 100;
      93 Best Local Similarity 100.0%; Pred. No. 1,07e-10;
      94 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
      95
      96 Db      84 EBYLILSARDVLAVSK 100
      97 1 |||||.....|||

```

Q? 1 EBYLLSARDVLAVWSK 17

ENTRY 4 S72818 #type complete
 TITLE heat shock protein chpA - Mycobacterium leprae
 ALTERNATE_NAMES chaperonin, 10K; protein Bl620_C3_227
 ORGANISM #formal_name Mycobacterium leprae
 DATE 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change

ACCESSIONS S72818
 REFERENCE S72580 Smith, D.R.; Robison, K.
 #authors Submitted to the EMBL Data Library, November 1993
 #description Mycobacterium leprae cosmid Bl620.
 #accession S72818
 #status preliminary
 #molecule_type DNA
 #residues 1-169 #label SMI
 #cross-references EXBL:U00015; NID:q466931; PIDN:AAQ43227.1;
 PID:q466939

GENETICS chpA
 CLASSIFICATION #superfamily chaperonin groES
 KEYWORDS heat shock; molecular chaperone; stress-induced protein
 SUMMARY #length 169 #molecular_weight 18097 #checksum 6011

Query Match 100.0%; Score 114; DB 2; Length 169;
 Best Local Similarity 100.0%; Pred. No. 1,07e-10;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

BL 153 EBYLLSARDVLAVWSK 169

Q? 1 EBYLLSARDVLAVWSK 17

ENTRY 5 S37565 #type complete
 TITLE chaperonin groES - Streptomyces coelicolor
 ORGANISM #formal_name Streptomyces coelicolor
 DATE 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change

ACCESSIONS S37565
 REFERENCE S37564 Carbone, A.M.; Kieser, K.; Hopwood, D.; Thompson, C.;
 Mazodier, P.
 #submission Submitted to the EMBL Data Library, September 1993
 #description Molecular characterization of two groEL genes in Streptomyces
 coelicolor A3(2).
 #accession S37565

GENETICS #status preliminary
 #molecule_type DNA
 #residues 1-102 #label DCC
 #cross-references EMBL:X75226; NID:q406595; PIDN:CAA53018.1;
 PID:q809755

ENTRY groES
 CLASSIFICATION #start_codon GTG
 #superfamily chaperonin groES
 KEYWORDS heat shock; molecular chaperone; stress-induced protein
 SUMMARY #length 102 #molecular_weight 10945 #checksum 4805

Query Match 92.1%; Score 105; DB 2; Length 102;
 Best Local Similarity 82.4%; Pred. No. 1.16e-08;
 Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

DB 86 EBYLV:SARDVLAVVEK 102

Q? 1 EBYLLSARDVLAVWSK 17

ENTRY 6 A41325 #type complete

TITLE heat shock protein 18 - Streptomyces albus
 ALTERNATE_NAMES heat shock protein groES homolog
 ORGANISM #formal_name Streptomyces albus
 DATE 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change

ACCESSIONS A41325
 REFERENCE A41325 Mazodier, P.; Guggenheim, G.; Davies, J.; Thompson, C.;
 #authors J. Bacteriol. (1991) 173:7362-7386
 #journal Characterization of the groEL-like genes in Streptomyces
 albus.

#cross-references MUID:92041639

#accession A41325

#molecule_type DNA

#residues 1-102 #label MAZ

#cross-references GB:M 57

GENETICS

#start_codon GTG

CLASSIFICATION #superfamily chaperonin groES

SUMMARY #length 102 #molecular_weight 10955 #checksum 4443

Query Match 93.4%; Score 103; DB 2; Length 102;
 Best Local Similarity 76.5%; Pred. No. 3.24e-08;
 Matches 13; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

DB 86 EBYLVLSARDVLAVVEK 102

Q? 1 EBYLLSARDVLAVWSK 17

ENTRY 7 A36721 #type complete
 TITLE groES protein - Synechococcus sp. (strain PCC 7942)
 ORGANISM #formal_name Synechococcus sp.
 DATE 12-Apr-1991 #sequence_revision 12-Apr-1991 #text_change

ACCESSIONS A36721
 REFERENCE A36721 Webb, R.; Reddy, K.J.; Sherman, L.A.;
 #authors J. Bacteriol. (1990) 172:5079-5088
 #journal Regulation and sequence of the Synechococcus sp. strain PCC
 7942 groES operon, encoding a cyanobacterial chaperonin.
 #cross-references MUID:9036856;

#accession A36721

#status preliminary

#molecule_type DNA

#residues 1-103 #label WEB

#cross-references GB:X58701; NID:q154519; PIDN:AAA27313.1; PID:q154520

CLASSIFICATION #superfamily chaperonin groES

SUMMARY #length 103 #molecular_weight 10742 #checksum 6554

Query Match 73.2%; Score 80; DB 2; Length 103;
 Best Local Similarity 50.0%; Pred. No. 2.39e-04;
 Matches 8; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

DB 88 DQVYLSEKDLAVVA 103

Q? 1 EBYLLSARDVLAVWS 16

ENTRY 8 BVYCGS #type complete
 TITLE chaperonin groES - Synechococcus sp. (strain PCC 6301)
 ALTERNATE_NAMES heat shock protein, 10K (hsp10)
 ORGANISM #formal_name Synechococcus sp.
 DATE 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change

ACCESSIONS S10836; S09619

REFERENCE SC7286

#authors Cozens, A.L.; Walker, J.E.

#journal J. Mol. Biol. (1987) 194:359-383

#title The organization and sequence of the genes for ATP synthase
 subunits in the cyanobacterium Synechococcus 6301. Support

 WIREP

 (TM)

Release 3.1A John F. Collins, BioComputing Research Unit.
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution Rights by Oxford Molecular Ltd

Search_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Jun 19 16:07:40 2000; Maspar time 3.67 seconds
 Tabular output not generated. 141.095 Million cell updates/sec

Title: XUS-09-142-524A-8
 Description: (1-17) from US09142524A.pep
 Perfect Score: 114
 Sequence: 1 REYLILSARDVLAVYSK 17

Scoring table: PAM 150
 Gap 15

Searched: 83857 seqs, 30454973 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: swiss-prot38
 1: swissprot

Statistics: Mean 29.618; Variance 37.117; scale 0.798

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query | Score | Match | Length | DB | ID | Description | Pred. No. |
|------------|-------|-------|-------|--------|------------|----|---------------------|-----------|
| 1 | 114 | 100.0 | 99 | 1 | CH10_MYCTU | 10 | KD CHAPERONIN (PROT | 9.22e-13 |
| 2 | 114 | 100.0 | 99 | 1 | CH10_MYCLE | 10 | KD CHAPERONIN (PROT | 9.22e-13 |
| 3 | 114 | 100.0 | 99 | 1 | CH10_MYCAV | 10 | KD CHAPERONIN (PROT | 9.22e-13 |
| 4 | 105 | 92.1 | 102 | 1 | CH10_STRO | 10 | KD CHAPERONIN (PROT | 1.94e-10 |
| 5 | 103 | 90.4 | 102 | 1 | CH10_STRAL | 10 | KD CHAPERONIN (PROT | 5.22e-10 |
| 6 | 80 | 76.2 | 100 | 1 | CH10_THER | 10 | KD CHAPERONIN (PROT | 2.13e-04 |
| 7 | 80 | 70.2 | 103 | 1 | CH10_SYNP7 | 10 | KD CHAPERONIN (PROT | 2.13e-04 |
| 8 | 80 | 70.2 | 103 | 1 | CH10_SYNP6 | 10 | KD CHAPERONIN (PROT | 2.13e-04 |
| 9 | 79 | 69.3 | 99 | 1 | CH10_MYCBU | 10 | KD CHAPERONIN (PROT | 3.58e-04 |
| 10 | 78 | 68.4 | 102 | 1 | CH10_SYNVU | 10 | KD CHAPERONIN (PROT | 6.00e-04 |
| 11 | 78 | 68.4 | 103 | 1 | CH10_CYPAP | 10 | KD CHAPERONIN (PROT | 5.00e-04 |
| 12 | 76 | 66.7 | 96 | 1 | CH10_LEPIN | 10 | KD CHAPERONIN (PROT | 1.67e-03 |
| 13 | 75 | 65.8 | 94 | 1 | CH10_BAGSU | 10 | KD CHAPERONIN (PROT | 2.77e-03 |
| 14 | 75 | 65.8 | 94 | 1 | CH10_BAGP3 | 10 | KD CHAPERONIN (PROT | 2.77e-03 |
| 15 | 74 | 64.9 | 94 | 1 | CH10_BACP3 | 10 | KD CHAPERONIN (PROT | 4.57e-03 |
| 16 | 74 | 64.9 | 104 | 1 | CH11_BRAJA | 10 | KD CHAPERONIN 1 (PR | 4.57e-03 |
| 17 | 73 | 64.0 | 94 | 1 | CH10_THERB | 10 | KD CHAPERONIN (PROT | 7.52e-03 |
| 18 | 71 | 62.3 | 94 | 1 | CH10_LACLA | 10 | KD CHAPERONIN (PROT | 2.01e-02 |
| 19 | 70 | 61.4 | 94 | 1 | CH10_STRAL | 10 | KD CHAPERONIN (PROT | 3.27e-02 |
| 20 | 70 | 61.4 | 96 | 1 | CH10_HAGDU | 10 | KD CHAPERONIN (PROT | 3.27e-02 |
| 21 | 59 | 60.5 | 96 | 1 | CH10_AC1PL | 10 | KD CHAPERONIN (PROT | 5.25e-02 |
| 22 | 59 | 60.5 | 102 | 1 | CH10_SYNP3 | 10 | KD CHAPERONIN (PROT | 5.25e-02 |
| 23 | 68 | 59.6 | 96 | 1 | CH10_ACYPS | 10 | KD CHAPERONIN (PROT | 8.52e-02 |

ALIGNMENTS

RESULT 1
 ID CH10_MYCTU STANDARD: PRI: 99 AA.
 AC P09621:
 DI 01-MAR-1989 (Rel. 10, Created;
 D: 15-FEB-2000 (Rel. 39, Last sequence update)
 DT 15-FEB-2000 (Rel. 39, Last annotation update)
 DE 10 KD CHAPERONIN (PROTEIN CPN10) (PROTEIN GROES) (BOG-A HEATSHOCK
 DE PROTEIN) (10 KD ANTIGEN).
 GN MOPB OR GROES OR CPN10 OR RV3418C OR MICV78 11.
 OS Mycobacterium tuberculosis.
 OC Bacteria: Firmicutes; Actinobacteria; Actinobacteridae; Mycobacteriaceae; Mycobacterium.
 CC Actinomycetales; Corynebacteriales; Mycobacteriales; Mycobacterium.
 RN 1:
 RP SEQUENC FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE: 89016584.
 RA Baird P.N., Hall L.M., Coates A.R.M.,
 RT "A major antigen from Mycobacterium tuberculosis which is homologous
 RT to the heat shock proteins GroES from E. coli and the hspA gene
 RT product of Coxiella burnetii".
 RL Nucleic Acids Res. 15:9047-9047(1988).
 RN 2:
 RP SEQUENC FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE: 89096443.
 RA Baird P.N., Hall L.M., Coates A.R.M.,
 RT "Cloning and sequence analysis of the 10 kDa antigen gene of
 RT Mycobacterium tuberculosis".
 RL J. Gen. Microbiol. 135:931-939(1989).
 RN 3:
 RP SEQUENC FROM N.A.
 RC STRAIN=ERMANN;
 RX MEDLINE: 89160258.
 RA Shinnick T.M., Plikaytis B.P., Hyshe A.D., van Landingham R.M.,
 RA Walker L.L.,
 RT "The Mycobacterium tuberculosis Bog-a protein has homology with the
 RT Escherichia coli GroES protein".
 RL Nucleic Acids Res. 17:1254-1254(1989).
 RN 4:
 RP SEQUENC FROM N.A.
 RC STRAIN=ERMANN;
 RX MEDLINE: 93219332.
 RA Kong I.H., Coates A.R.M., Butcher P.D., Hickman C.J., Shinnick T.M.,
 RT "Mycobacterium tuberculosis expresses two chaperonin-60 homologs".
 RL Proc. Natl. Acad. Sci. U.S.A. 90:2508-2512(1993).
 RN 5:
 RP SEQUENC FROM N.A.

RC STRAIN: H37RV;
 RX MEDLINE: 98295987.
 RA Cole S., Broese R., Parkhill J., Garnier J., Churcher C., Harris D.,
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Felwell T., Gentles S., Hill N., Holroyd S.,
 RA Horvath S., Jazels K., Krogh A., McLean J., Moult S., Murphy L.,
 RA Oliver S., Osborne J., Quail A.A., Rajandream K.A., Rogers J.,
 RA Rucker S., Seeger K., Skellern S., Squares S., Squares R., Sulston J.E.,
 RA Taylor K., Whitehead S., Barrett B.G.,
 RI "Deciphering the biology of Mycobacterium tuberculosis from the
 RI complete genome sequence."
 RI Nature 393:537-544 (1998).
 RN 16.
 RP SEQUENCE OF 1-15.
 RX MEDLINE: 92174646.
 RA Barnes P.F., Mehra V., Riviere B., Ford S.J., Brennan P.J.,
 RA Voegtline M.S., Minden P., Houghton R.A., Bloom B.R., Modlin R.L.,
 RI "Immunoreactivity of a 10-kDa antigen of Mycobacterium tuberculosis."
 RI J. Immunol. 148:1835-1843 (1992).
 CC 1- FUNCTION: BINDS TO CPN10 IN THE PRESENCE OF MG-ATP AND SUPPRESSES
 CC THE ATPASE ACTIVITY OF THE LATTER.
 CC 1- SUBUNIT: HEPTAMER OF 7 SUBUNITS ARRANGED IN A RING
 CC (BY SIMILARITY).
 CC 1- SIMILARITY: BELONGS TO THE GROES CHAPERONIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X63350; CAA42902.1;
 CC EMBL: X12598; CAA31111.1;
 CC EMBL: M25258; AAZ25340.1;
 CC EMBL: X13739; CAA32003.1;
 CC EMBL: Z71145; CAA31005.1;
 CC EMBL: S02743; BVMV5A.
 CC PIR: A37165; A37166.
 CC PIR: A47292; A47292.
 CC PRINTS: PR00297; CHAPERONIN10.
 CC PROSITE: PS00681; CHAPERONIN10-CPN10: 1.
 CC PFAM: PF00166; CPN10: 1.
 CC TUBERCULIST: T33418c; 1.
 CC Chaperone: Antigen: Heat shock.
 CC INIT_MET 0
 CC SEQUENCE 99 AA: 10673 MW: 100228575C5F59A7 CP-64;
 Query Match 100.0%; Score 114; DB 1; Length 99;
 Best Local Similarity 100.0%; Pred. No. 9,22e-13;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DI 83 EYLLSARDVLAVVSK 99
 QY 1 EYLLSARDVLAVVSK 17
 RESULT 2
 ID CH10_MYCAV STANDARD: PRT: 99 AA.
 AC P24331
 DI 01-MAR-1992 (Rel. 21; Created)
 DI 01-MAR-1992 (Rel. 21; Last sequence update)
 DI 15-DEC-1998 (Rel. 37; Last annotation update)
 DE 10 KD CHAPERONIN (PROTEIN CPN10) (PROTEIN GROES) (10 KD ANTIGEN).
 GN MOPB OR GROES OR CHPA OR B1620_C3_227 OR B229_C3_247.
 OS Mycobacterium leprae.
 CC Bacteria; Firmicutes.
 CC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacteri-
 CC Bacteria; Firmicutes.
 RN RN
 RE SEQUENCE FROM N.A., AND SEQUENCE OF 1-24.
 RX MEDLINE: 92113469.

RA Mehra V., Bloom B.R., Bajardi A.C., Grisso A.C., Sieling P.A.,
 RA Alland D., Convit J., Fan X., Hunter S.W., Brennan P.J., Red T.H.,
 RA Modlin R.L.,
 RI "A major T cell antigen of Mycobacterium leprae is a 10-kD heat-shock
 RI cognate protein."
 RI J. Exp. Med. 175:275-284 (1992).
 RL 2.
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 92374850.
 RA de Wit T.F.R., Bekke S., Osleid A., Mink J.L., Hermans P.W.M.
 RA van Soelingen D., Driifhout J., Schoeninch R., Janson A.A.K.,
 RA Thole J.E.R.,
 RI "Mycobacteria contain two groEL genes: the second Mycobacterium
 RI leprae groEL gene is arranged in an operon with groES."
 RI Mol. Microbiol. 6:1995-2007 (1992).
 RN 13.
 RP SEQUENCE FROM N.A.
 RA Smith D.R., Robison K.,
 RI Submitted (MAR-1994) to the EMBL/Genbank/DBS databases.
 RN 4.
 RX X-RAY CRYSTALLOGRAPHY (3.5 ANGSTROMS).
 RA Made S.C., Mehra V., Bloom B.R., Hol W.S.J.,
 RI "Structure of the heat shock protein chaperonin-10 of Mycobacterium
 RI leprae."
 RI Science 271:203-207 (1996).
 RN 5.
 RP ERRATUM.
 RA Made S.C., Mehra V., Bloom B.R., Hol W.S.J.,
 RI Science 271:1655-1655 (1996).
 CC 1- FUNCTION: BINDS TO CPN10 IN THE PRESENCE OF MG-ATP AND SUPPRESSES
 CC THE ATPASE ACTIVITY OF THE LATTER.
 CC 1- SUBUNIT: HEPTAMER OF 7 SUBUNITS ARRANGED IN A RING.
 CC 1- SIMILARITY: BELONGS TO THE GROES CHAPERONIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X59413; NOT ANNOTATED CDS.
 CC EMBL: Z11665; NOT ANNOTATED CDS.
 CC EMBL: C06015; AAC43227.1;
 CC EMBL: U00020; AAL731.1;
 CC PIR: ZC0556; ZC0556.
 CC PIR: Z23180; Z23180.
 CC PCB: ILEP; ILEP-47.
 CC PRINTS: PR00397; CHAPERONIN10-CPN10: 1.
 CC PROSITE: PS00681; CHAPERONIN10-CPN10: 1.
 CC PFAM: PF00166; CPN10: 1.
 CC Chaperone: Antigen: Heat shock; 4D-structure.
 CC INIT_MET 0
 CC SEQUENCE 99 AA: 10669 MW: 60288983F7F5EB CRC64;
 Query Match 100.0%; Score 114; DB 1; Length 99;
 Best Local Similarity 100.0%; Pred. No. 9,22e-13;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DI 83 EYLLSARDVLAVVSK 99
 QY 1 EYLLSARDVLAVVSK 17
 RESULT 3
 ID CH10_MYCAV STANDARD: PRT: 99 AA.
 AC O84017;
 DI 15-FEB-2000 (Rel. 39; Created)
 DI 15-FEB-2000 (Rel. 39; Last sequence update)
 DI 15-FEB-2000 (Rel. 39; Last annotation update)
 DE 10 KD CHAPERONIN (PROTEIN CPN10) (PROTEIN GROES) (10 KD ANTIGEN).

GN MOPR OR GROES.
 CS Mycobacterium avium, and Mycobacterium paratuberculosis.
 CC Bacteria: Firmicutes: Actinobacteria: Actinobacteridae;
 CC Actinomycetales: Corynebacteriineae: Mycobacteriaceae: Mycobacterium.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES-M.AVIUM: STRAIN-485 TYPE 211
 RA Cresti R., Pietrobono R., Rattorini L., Orefici G.,
 RI "Sequence of the groEL operon of Mycobacterium avium comprising the
 RI gene encoding the cpn60 protein and a portion of the gene encoding the
 RI cpn60-1 protein."
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES-M.PARATUBERCULOSIS: M.AVIUM: STRAIN-ATCC 19698, ATCC 25291;
 RA Cobb A.J., Frothingham R.;
 RI "The GroES antigens of Mycobacterium avium and Mycobacterium
 RI paratuberculosis."
 RL Vet. Microbiol. 0:0-0(1999).
 CC -1- FUNCTION: BINDS TO CPN60 IN THE PRESENCE OF MG-ATP AND SUPPRESS
 CC THE ATPASE ACTIVITY OF THE LATTER (BY SIMILARITY).
 CC -1- SUBUNIT: HEPTAMER OF 7 SUBUNITS ARRANGED IN A RING (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE GROES CHAPERONIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: AF079544; AAC31921.1;
 DR EMBL: AF071829; AAD23277.1;
 DR EMBL: AF071828; AAD23276.1;
 DR PFAM: PF00166; cpn10; 1;
 DR PRINTS: PR00267; CHAPERONIN10;
 DR PROSITE: PS00681; CHAPERONINS_CPN10; 1;
 KW Chaperone; Antigen; Heat shock;
 FI INIT_MET C 0 BY SIMILARITY.
 SV SEQUENCE 99 AA: 10517 MW: 6804437405310C2 CRC64:
 Query Match 100.0% Score 114; DB 1; Length 99;
 Best Local Similarity 100.0%; Pred. No. 9,22e-13;
 Matches 27: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 83 EBYLVLSARDVLAVSK 49
 QY 1 EBYLVLSARDVLAVSK 17
 |||||
 RESULT 4
 ID CH10_STRCO STANDARD: PRI: 102 AA.
 AC 94072;
 DT 31-FEB-1995 (Rel. 31; Created);
 DT 01-FEB-1995 (Rel. 31; Last sequence update);
 DT 15-DEC-1996 (Rel. 37; Last annotation update)
 DE 10 KD CHAPERONIN (PROTEIN CPN10) (PROTEIN GROES).
 GN GROES OR SC634.39
 CC Streptomyces coelicolor, and Streptomyces lividans.
 CC Bacteria: Firmicutes: Actinobacteria: Actinobacteridae;
 CC Actinomycetales: Streptomycineae: Streptomycetaceae: Streptomyces.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES-S.COELICOLOR: STRAIN-A3(2) / J1501;
 RA MEDLINE: 94294177.
 RA Duchene A.M., Kieser H.M., Hopwood D.A., Thompson C.J., Mazodier P.;
 RI "Characterization of two groEL genes in Streptomyces coelicolor:
 RI A3(2)."
 RL Gene 144:97-101(1994).
 RN [2]
 RP SEQUENCE FROM N.A.

RC SPECIES-S.COELICOLOR: STRAIN-A3(2);
 RA Saunders D., Harris D., Parkhill J., Barrrell B.G., Ralmandream M.A.;
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES-S.LIVIDANS;
 RX MEDLINE: 98048481;
 RA de Leon P., Marco S., Islegas C., Marina A., Carrascosa J.L.;
 RI "Streptomyces lividans groES, groEL1 and groEL2 genes."
 RL Microbiology 143:3563-3571(1997).
 CC -1- FUNCTION: BINDS TO CPN60 IN THE PRESENCE OF MG-ATP AND SUPPRESSES
 CC THE ATPASE ACTIVITY OF THE LATTER.
 CC -1- SUBUNIT: HEPTAMER OF 7 SUBUNITS ARRANGED IN A RING
 CC (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE GROES CHAPERONIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: X75206; CAA53018.1;
 DR EMBL: ALC31317; CAA20417.1;
 DR EMBL: X95970; CAA65224.1;
 DR PIR: S37565; S37565;
 DR PRINTS: PR00297; CHAPERONIN10;
 DR PROSITE: PS00681; CHAPERONINS_CPN10; 1;
 DR PFAM: PF00166; cpn10; 1;
 KW Chaperone.
 SV SEQUENCE 102 AA: 10946 MW: 60F9FA68014AC8CC CRC64:
 Query Match 92.1% Score 105; DB 1; Length 102;
 Best Local Similarity 82.4%; Pred. No. 1.94e-19;
 Matches 14: Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 DB 86 EBYLVLSARDVLAVSK 102
 QY 1 EBYLVLSARDVLAVSK 17
 |||||
 RESULT 5
 ID CH10_STRAL STANDARD: PRI: 102 AA.
 AC Q00769;
 DT 01-APR-1993 (Rel. 25; Created);
 DT 01-APR-1993 (Rel. 25; Last sequence update);
 DT 01-FEB-1996 (Rel. 33; Last annotation update);
 DE 10 KD CHAPERONIN (PROTEIN CPN10) (PROTEIN GROES).
 GN GROES
 CC Streptomyces albus J.
 CC Bacteria: Firmicutes: Actinobacteria: Actinobacteridae;
 CC Actinomycetales: Streptomycineae: Streptomycetaceae: Streptomyces.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 92041639.
 RA Mazodier P., Gahleitner G., Davies J., Thompson C.J.;
 RI "Characterization of the groEL-like genes in Streptomyces albus J."
 RL Bacteriol. 173:7382-7386(1994).
 CC -1- FUNCTION: BINDS TO CPN60 IN THE PRESENCE OF MG-ATP AND SUPPRESSES
 CC THE ATPASE ACTIVITY OF THE LATTER.
 CC -1- SUBUNIT: HEPTAMER OF 7 SUBUNITS ARRANGED IN A RING
 CC (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE GROES CHAPERONIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

```

CC or send an email to license@lsb-sib.ch).
CC
CC
CC
CC EMBL: M76657; AAA26752.1;
CC PRINTS: PR00297; CHAPERONIN10.
CC PROSITE: PS00681; CHAPERONINS_CPN10; 1.
CC PFAM: PF00166; cpn10; 1.
CC GROES: GR HSP10.
CC CHAPERONE: 102 AA; 10987 MW; 5780F5554CFB283 CRC64;
CC SEQUENCE: 102 AA; 10987 MW; 5780F5554CFB283 CRC64;

Query Match 90.4% Score 103; CB 1; Length 102;
Best Local Similarity 76.5% Pred. No. 5.21e-04;
Matches 13; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

DB 85 EYVILSARDVLAV 16
QY 1 EYVILSARDVLAV 16

RESULT 5
ID CH10_SYNPF STANDARD: PRT: 100 AA.
AC P2280;
DI 01-AUG-1991 (Rel. 19, Created)
DI 01-AUG-1991 (Rel. 19, Last sequence update)
DI 01-FEB-1996 (Rel. 33, Last annotation update)
DE 10 KD CHAPERONIN (PROTEIN CPN10) (PROTEIN GROES).
GN GROES.
OS Thermus aquaticus (subsp. thermophilus).
OC Bacteria; Thermus/Delfinococcus group; thermus group; Thermus.
RN 11.
RP SEQUENCE FROM N.A.
RA MEDLINE: 95044154.
RA Amada K., Yoda M., Oda K., Endo I., Ishii N., Tanuchi H.,
RA Yoshida M.;
RA "Molecular cloning, expression, and characterization of chaperonin-60
RA and chaperonin-10 from a thermophilic bacterium, Thermus thermophilus
RA HBA."
RA J. Biochem. 116:347-354 (1995).
RA [2].
RP SEQUENCE FROM N.A.
RA MEDLINE: 92042183.
RA Tanuchi H., Konishi J., Ishii N., Yoshida M.;
RA "A chaperonin from a thermophilic bacterium, Thermus thermophilus,
RA that controls refoldings of several thermophilic enzymes."
RA J. Biol. Chem. 266:22411-22419 (1991).
CC -- FUNCTION: BINDS TO CPN60 IN THE PRESENCE OF MG-ATP AND SUPPRESSES
CC THE ATPASE ACTIVITY OF THE LATTER.
CC -- SUBUNIT: HEPTAMER OF 7 SUBUNITS ARRANGED IN A RING
CC (BY SIMILARITY).
CC -- SIMILARITY: BELONGS TO THE GROES CHAPERONIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC
CC EMBL: M58751; AAA27313.1;
CC PIR: A36721; A36721.
CC PRINTS: PR00297; CHAPERONIN10.
CC PROSITE: PS00681; CHAPERONINS_CPN10; 1.
CC PFAM: PF00166; cpn10; 1.
CC CHAPERONE: 103 AA; 10742 MW; 89107145F515452E CRC64.
CC SEQUENCE: 103 AA; 10742 MW; 89107145F515452E CRC64.

Query Match 76.2% Score 103; CB 1; Length 103;
Best Local Similarity 55.0% Pred. No. 2.14e-04;
Matches 8; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

DB 88 DDYVLLSEKDLAV 103
QY 1 EYVILSARDVLAV 16

RESULT 6
ID CH10_SYNPF STANDARD: PRT: 103 AA.
AC P07895;
DI 01-AUG-1988 (Rel. 08, Created)
DI 01-AUG-1988 (Rel. 08, Last sequence update)
DI 01-FEB-1996 (Rel. 33, Last annotation update)
DE 10 KD CHAPERONIN (PROTEIN PN10) (PROTEIN GROES).
GN GROES.
OS Synecococcus sp. (strain PC 5301).
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
RN 11.
RP SEQUENCE FROM N.A.
RA MEDLINE: 87311713.
RA

```

```

FT CONFLICT 83 86 13EE -> APHRT (IN REF. 27).
SQ SEQUENCE 100 AA; 10965 MW; 452F30C9CCBC0FA CRC64;

Query Match 70.2% Score 103; CB 1; Length 100;
Best Local Similarity 73.3% Pred. No. 2.13e-04;
Matches 13; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

DB 85 EYVILSARDVLAV 99
QY 1 EYVILSARDVLAV 15

RESULT 7
ID CH10_SYNPF STANDARD: PRT: 103 AA.
AC P2280;
DI 01-AUG-1991 (Rel. 19, Created)
DI 01-AUG-1991 (Rel. 19, Last sequence update)
DI 01-FEB-1996 (Rel. 33, Last annotation update)
DE 10 KD CHAPERONIN (PROTEIN CPN10) (PROTEIN GROES).
GN GROES.
OS Synecococcus sp. (strain PC 5301) (Anacystis nidulans 542).
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
RN 11.
RP SEQUENCE FROM N.A.
RA MEDLINE: 90368561.
RA Webb R., Reddy K.S., Sherman L.A.;
RA "Regulation and sequence of the Synecococcus sp. strain PC 5301
RA groESL operon, encoding a cyanobacterial chaperonin."
RA J. Bacteriol. 172:5079-5084 (1990).
CC -- FUNCTION: BINDS TO CPN60 IN THE PRESENCE OF MG-ATP AND SUPPRESSES
CC THE ATPASE ACTIVITY OF THE LATTER.
CC -- SUBUNIT: HEPTAMER OF 7 SUBUNITS ARRANGED IN A RING
CC (BY SIMILARITY).
CC -- SIMILARITY: BELONGS TO THE GROES CHAPERONIN FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb-sib.ch).
CC
CC EMBL: M58751; AAA27313.1;
CC PIR: A36721; A36721.
CC PRINTS: PR00297; CHAPERONIN10.
CC PROSITE: PS00681; CHAPERONINS_CPN10; 1.
CC PFAM: PF00166; cpn10; 1.
CC CHAPERONE: 103 AA; 10742 MW; 89107145F515452E CRC64.
CC SEQUENCE: 103 AA; 10742 MW; 89107145F515452E CRC64.

Query Match 76.2% Score 103; CB 1; Length 103;
Best Local Similarity 55.0% Pred. No. 2.14e-04;
Matches 8; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

DB 88 DDYVLLSEKDLAV 103
QY 1 EYVILSARDVLAV 16

RESULT 8
ID CH10_SYNPF STANDARD: PRT: 103 AA.
AC P07895;
DI 01-AUG-1988 (Rel. 08, Created)
DI 01-AUG-1988 (Rel. 08, Last sequence update)
DI 01-FEB-1996 (Rel. 33, Last annotation update)
DE 10 KD CHAPERONIN (PROTEIN PN10) (PROTEIN GROES).
GN GROES.
OS Synecococcus sp. (strain PC 5301).
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
RN 11.
RP SEQUENCE FROM N.A.
RA MEDLINE: 87311713.
RA

```



```

RESULT 11
ID CHIC_CYPARA STANDARD: PSI: 103 AA.
AC Q37762
DT 01-NOV-1997 (Rel. 35, Created)
DI 01-NOV-1997 (Rel. 35, Last sequence update)
DI 15-DEC-1998 (Rel. 37, Last annotation update)
DE 10 KD CHAPERONIN (PROTEIN CPN10) (PROTEIN GROES)
GN GROES-A AND GROES-B
OS Cyanochoira parvula
OC Cyanelle
OC Eukaryota: Glomerisophyceae, Cyanothraaceae: Cyanothra.
RN [1]
RF SEQUENCE FROM N.A.
RC STRAIN LR555 / PRINGSHEIM
RA Stirewalt V.L., Michalowski C.B., Lottelhard W., Bohnert H.J.,
RA Bryant J.A.
RL Submitted (OCT-1975) to the EMBL/GenBank/DBCC databases.
CC -1- FUNCTION: BINDS TO CPN60 IN THE PRESENCE OF MG-ATP AND SUPPRESSES
CC THE ATPASE ACTIVITY OF THE LATTER (BY SIMILARITY).
CC -1- SUBUNIT: HEPTAMER OF 7 SUBUNITS ARRANGED IN A RING
CC (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE GROES CHAPERONIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U30821; AAA81731.1
CC EMBL: U30821; AAA82971.1
CC PRINTS: PRO0297; CHAPERONIN10
CC PROSITE: PS00681; CHAPERONIN10
CC PRAM: PF00166; CPN10
CC Chapteronin, Cyanelle.
KW CHAPERONIN
SQ SEQUENCE 103 AA: 11249 MW: 17664790CA90520 CR *64:
.....
Query Match 68.48: Score 78; DE 11; Length 103;
Best Local Similarity 50.00; Pred. No. 4,00e-04;
Matches 8; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

DB 88 EYLLSARDVLAVVSK 103
ID EYLLSARDVLAVVSK 103
AC EYLLSARDVLAVVSK 103
DT 01-JUN-1994 (Rel. 29, Created)
DI 01-JUN-1994 (Rel. 29, Last sequence update)
DI 15-DEC-1998 (Rel. 37, Last annotation update)
DE 10 KD CHAPERONIN (PROTEIN CPN10) (PROTEIN GROES) (HEAT SHOCK 10 KD
DE PROTEIN)
GN GROES OR HSP10.
OS Leptospira interrogans.
OC Bacteria: Spirochaetales: Leptospiraceae: Leptospira.
RN [1]
RF SEQUENCE FROM N.A.
RC STRAIN-SEROVAR COPENHAGEN1 / WIJNBORG
RA Ballard S.A., Seders R.P., Bloumick-Plyum N.M., Fyfe J.A.M.,
RA Faine S., Adler B.
RT "Molecular analysis of the hsp (groE) operon of Leptospira
RT interrogans serovar Copenhagen1."
RL Mol. Microbiol. 8:739-751(1993).
RN [2]
RF SEQUENCE FROM N.A.
RC STRAIN-SEROVAR HAT / HY-1;
RA Kim M.J., Ahn E.Y.;

```

```

Submitted (NOV-1997) to the EMBL/GenBank/DBCC databases.
CC -1- FUNCTION: BINDS TO CPN60 IN THE PRESENCE OF MG-ATP AND SUPPRESSES
CC THE ATPASE ACTIVITY OF THE LATTER.
CC -1- SUBUNIT: HEPTAMER OF 7 SUBUNITS ARRANGED IN A RING
CC (BY SIMILARITY).
CC -1- INDUCTION: BY HEAT SHOCK.
CC -1- SIMILARITY: BELONGS TO THE GROES CHAPERONIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U34502; AAA71931.1
CC EMBL: AF032910; AAB95964.1
CC PRINTS: S34937; S34937
CC PROSITE: PS00297; CHAPERONIN10
CC PROSITE: PS00681; CHAPERONIN10
CC PRAM: PF00166; CPN10
CC Chapterone, Heat shock.
KW CHAPERONE
SQ SEQUENCE 96 AA: 10562 MW: 925507989F1467 CR *64:
.....
Query Match 66.78: Score 76; DE 11; Length 96;
Best Local Similarity 52.58; Pred. No. 1.57e-03;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

DB 81 EYLLSARDVLAVVSK 96
ID EYLLSARDVLAVVSK 96
AC EYLLSARDVLAVVSK 96
DT 01-JUN-1994 (Rel. 24, Created)
DI 01-JUN-1994 (Rel. 26, Last sequence update)
DI 15-DEC-1998 (Rel. 37, Last annotation update)
DE 10 KD CHAPERONIN (PROTEIN CPN10) (PROTEIN GROES)
GN GROES.
OS Bacillus subtilis.
OC Bacteria: Firmicutes: Bacillus/Clostridium group:
OC Bacillus/Staphylococcus group: Bacillus.
RN [1]
RF SEQUENCE FROM N.A.
RC STRAIN: 92283754
RA Schmidt A., Schaefer M., Winkler M., Bockel M., Schmeckel N.,
RA "Cloning, sequencing, mapping, and functional analysis of the
RA groES operon from Bacillus subtilis."
RA J. Bacteriol. 174:1997-1999(1992).
RN [2]
RF SEQUENCE FROM N.A.
RC STRAIN:168;
RA MEDLINE: 92283753.
RA Li M., Wong S.L.
RT "Cloning and characterization of the groES operon from Bacillus
RT subtilis."
RN [3]
RF Bacterio.. 174:3981-3992(1992).
RN [3]
RF SEQUENCE FROM N.A.
RC STRAIN 168 / MARGURG;
RA MEDLINE: 93129852.
RA Tozawa Y., Yoshikawa H., Kawamura F., Itaya M., Takahashi H.
RT "Isolation and characterization of the groES and groEL genes of
RT Bacillus subtilis Marburg."
RA Biosci. Biotechnol. Biochem. 56:1995-2002(1992).
RN [4]
RF SEQUENCE FROM N.A.
RC STRAIN:168 / MARGURG;
RA MEDLINE: 93746038.
RA Sadale Y., Yata K., Fujita M., Sazai H., Itaya M., Kasahara Y.

```


CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

DR EMBL: S57424: CAB31564.1: ALI_SEQ.

DR PIR: JQ1194: JQ1194.

DR PIR: JCI479: JCI479.

DR PRINTS: PR00297: CHAPERONIN10.

DR PROSITE: PS00681: CHAPERONINS_cpnl0: 1.

DR PFAM: PF00166: cpnl0: 1.

DR Chaperone.

SC SEQUENCE 94 AA: 10340 MW: 24944954AE0B457C CR664:

Query Match 54.9% Score 74: DB 1: Length 94:

Best Local Similarity 64.3% Pred. No. 4.57e-03:

Matches 9: Conservative 2: Mismatches 3: Indels 0: Gaps 0:

DR 80 EYLILRESILAV: 94

QY 1111 1111 1

2 EYLILFARGVLAV: 16

Search completed: Mon Jun 19 16:07:46 2000

Job time: 6 secs.

 WAREHOUSE (TM)

Release 3.1A John F. Collins, BioComputing Research Unit.

Copyright (c) 1993-1998 University of Edinburgh, U.K.

Distribution rights by Oxford Molecular Ltd

MFsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Jun 19 16:07:11 2000 Maspar time 8.85 Seconds

Tabular output not generated. 133.146 Million cell updates/sec

Title: US-09-142-524A-8

Description: (1-17) from US09142524A.pep

Perfect Score: 114

Sequence: 1 EEVL12SARCVIAVYSK 17

Scoring table: BAW 150

Gap 15

Searched: 225878 seqs. 59334122 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: spiremb12

1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human
 5:sp-invertebrate 6:sp-mammal 7:sp-mhc 8:sp-organelle
 9:sp-phage 10:sp-plant 11:sp-protein 12:sp-unclassified
 13:sp-vertebrate 14:sp-virus

Statistics: Mean 28.205; Variance 39.500; scale 0.713

Pred. No. is the number of results predicted by chance to have a
 Score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description | Pred. No. |
|------------|-------|-------------|--------|-------|-------------------------------|-----------|
| 1 | 114 | 100.0 | 100 | 2 | 086017 10 KD CHAPERONIN (PROT | 2.76e-11 |
| 2 | 76 | 65.7 | 103 | 2 | Q9X205 10 KD CHAPERONIN (PROT | 8.15e-03 |
| 3 | 73 | 64.0 | 88 | 2 | Q50324 GROES (FRAGMENT) | 3.27e-02 |
| 4 | 70 | 61.4 | 100 | 2 | Q9XCBC 10 KD CHAPERONIN (PROT | 1.27e-01 |
| 5 | 68 | 59.6 | 94 | 2 | Q86323 10 KD CHAPERONIN (PROT | 3.10e-01 |
| 6 | 67 | 58.8 | 93 | 2 | G32846 10 KD CHAPERONIN (PROT | 4.81e-01 |
| 7 | 67 | 58.8 | 122 | 2 | G27942 GROES | 4.81e-01 |
| 8 | 67 | 58.8 | 304 | 13 | Q9W534 MRNA, PARTIAL CDS, SIM | 4.81e-01 |
| 9 | 67 | 59.8 | 312 | 13 | Q9W525 XIMPACT | 4.81e-01 |
| 10 | 66 | 57.9 | 455 | 4 | Q94899 KIAA0795 PROTEIN (FRAG | 7.43e-01 |
| 11 | 65 | 57.0 | 88 | 2 | Q66201 SIMILAR TO GROES PROTE | 1.14e-00 |
| 12 | 65 | 57.0 | 88 | 2 | Q66189 SIMILAR TO GROES PROTE | 1.14e-00 |
| 13 | 65 | 57.0 | 88 | 2 | Q66211 SIMILAR TO GROES PROTE | 1.14e-00 |
| 14 | 65 | 57.0 | 88 | 2 | Q66193 SIMILAR TO GROES PROTE | 1.14e-00 |
| 15 | 65 | 57.0 | 88 | 2 | Q66219 SIMILAR TO GROES PROTE | 1.14e-00 |
| 16 | 65 | 57.0 | 88 | 2 | Q66203 SIMILAR TO GROES PROTE | 1.14e-00 |
| 17 | 65 | 57.0 | 88 | 2 | Q66215 SIMILAR TO GROES PROTE | 1.14e-00 |
| 18 | 65 | 57.0 | 88 | 2 | Q66205 SIMILAR TO GROES PROTE | 1.14e-00 |
| 19 | 65 | 57.0 | 88 | 2 | Q66195 SIMILAR TO GROES PROTE | 1.14e-00 |
| 20 | 65 | 57.0 | 88 | 2 | Q66193 SIMILAR TO GROES PROTE | 1.14e-00 |

| | | | | | | |
|----|----|------|-----|----|--------------------------------|----------|
| 21 | 65 | 57.0 | 88 | 2 | Q66207 SIMILAR TO GROES PROTE | 1.14e-00 |
| 22 | 65 | 57.0 | 88 | 2 | Q66213 SIMILAR TO GROES PROTE | 1.14e-00 |
| 23 | 65 | 57.0 | 88 | 2 | Q66209 SIMILAR TO GROES PROTE | 1.14e-00 |
| 24 | 65 | 57.0 | 88 | 2 | Q66217 SIMILAR TO GROES PROTE | 1.14e-00 |
| 25 | 65 | 57.0 | 88 | 2 | Q66211 SIMILAR TO GROES PROTE | 1.14e-00 |
| 26 | 65 | 57.0 | 154 | 2 | Q66213 HYPOTHETICAL 17.8 KL I | 1.14e-00 |
| 27 | 64 | 56.1 | 89 | 2 | Q66213 GROES (FRAGMENT) | 1.14e-00 |
| 28 | 64 | 56.1 | 97 | 2 | Q66213 10 KD CHAPERONIN (PROT | 1.14e-00 |
| 29 | 63 | 55.3 | 88 | 2 | Q66217 SIMILAR TO GROES PROTE | 2.47e-00 |
| 30 | 63 | 55.3 | 88 | 2 | Q66221 SIMILAR TO GROES PROTE | 2.47e-00 |
| 31 | 63 | 55.3 | 95 | 2 | Q51831 10 KD CHAPERONIN (PROT | 2.47e-00 |
| 32 | 63 | 55.3 | 261 | 3 | Q42628 32 KDA-CELL WALL SYMB | 2.47e-00 |
| 33 | 63 | 55.3 | 330 | 8 | Q92203 NADH-UBIQUINONE OXIDOR | 2.47e-00 |
| 34 | 63 | 55.3 | 445 | 2 | Q66213 NEGATIVE ALPHATIC AM | 2.47e-00 |
| 35 | 63 | 55.3 | 629 | 10 | Q66213 PUTATIVE ACIDIC DEHYD | 2.47e-00 |
| 36 | 63 | 55.3 | 302 | 4 | Q66213 X-LIKE 1 PROTEIN | 2.47e-00 |
| 37 | 62 | 54.4 | 305 | 1 | Q66213 HYPOTHETICAL 14.8 KD P | 2.47e-00 |
| 38 | 61 | 53.5 | 114 | 5 | Q66213 HEMOLYSIN | 5.14e-00 |
| 39 | 61 | 53.5 | 206 | 2 | Q66213 THYMIDYLATE KINASE | 5.14e-00 |
| 40 | 61 | 53.5 | 242 | 5 | Q66213 ERCC1-LIKE EXCISION RE | 5.14e-00 |
| 41 | 61 | 53.5 | 300 | 5 | Q66213 SERINE/THREONINE PROT | 5.14e-00 |
| 42 | 61 | 53.5 | 305 | 5 | Q66213 CRK-RELATED PROTEIN K1 | 5.14e-00 |
| 43 | 61 | 53.5 | 370 | 5 | Q66213 CUSC8.2 PROTEIN | 5.14e-00 |
| 44 | 61 | 53.5 | 387 | 1 | Q66213 CRP C02004 | 5.14e-00 |
| 45 | 61 | 53.5 | 476 | 2 | Q66213 PUTATIVE AMINO TRANSFER | 5.14e-00 |

ALIGNMENTS

RESULT 1
 ID 086017 PRELIMINARY; PRT: 100 AA.
 AC 086017;
 DT 01-NOV-1998 (TREMURel. 08. Created)
 DT 01-NOV-1998 (TREMURel. 08. Last sequence update)
 DT 01-NOV-1998 (TREMURel. 12. Last annotation update)
 DE 10 KD CHAPERONIN (PROTEIN CPN10) (PROTEIN GROES).
 GN GROES.
 OS Mycobacterium avium, and Mycobacterium paratuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 RN [1]
 PP SEQUENCE FROM N.A.
 RC STRAIN-495 TYPE 21.
 RA CREIG R., PETROBOGO R., PATRINI L., GREFFICI G.
 RT "Sequence of the GroES operon of Mycobacterium avium comprising the
 RI gene encoding the cpn10 protein and a portion of the gene encoding the
 RI cpn60-1 protein."
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 PP SEQUENCE FROM N.A.
 RC SPECIES: PARATUBERCULOSIS, STRAIN-495 TYPE 21.
 RA COBB A.J., PROTHINGHAM R.
 RT "The GroES antigens of Mycobacterium avium and Mycobacterium
 RI paratuberculosis."
 RL Vet. Microbiol. 0:0-0(1999).
 CC -1- FUNCTION: BINDS TO CPN60 IN THE PRESENCE OF MG-AIP AND SUPPRESSES
 CC THE ATPASE ACTIVITY OF THE LATTER (BY SIMILARITY).
 CC -1- SUBUNIT: HEPTAMER OF 7 SUBUNITS ARRANGED IN A RING (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE GROES CHAPERONIN FAMILY.
 DR EMBL: AF079544; AAC31921.1;
 DR EMBL: AF071829; AAD23277.1;
 DR EMBL: AF071828; AAD23276.1;
 DR PFM: PFM0166; cpn10; 1.
 DR PROSITE: PS00681; CHAPERONINS_CPN10; 1.
 DR PRINTS: PR00297; CHAPERON N10.
 KW Chaperone; Heat shock.
 SQ SEQUENCE 100 AA: 10748 M 0C4954DA CRC32;
 Query Match 100.0% score 114; DB 2; Length 100;
 Best Local Similarity 100.0%; Pred.No. 2.76e-11;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Q: 1 EYLILSARDVLAV 14

RESULT 6
ID Q32846 PRELIMINARY: PRT: 93 AA.
AC Q32846
DI 01-JAN-1998 (TRENBLrel. 05, Created)
DI 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DI 01-NOV-1999 (TRENBLrel. 12, Last annotation update)
DI 10 KD CHAPERONIN (PROTEIN CPN10) (PROTEIN GROES)
GN GROES
OS Lactobacillus zeae.
OT Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
OL Lactobacillus.
RN [1]
RP SEQUENCE FROM N.A.
R1 STRAIN-1025 (PLASMID-FREE DERIVATIVE - ATCC3939)
R2 MURPHY C.M., CHASSY B.M.
R3 Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
R4 FUNCTION: BINDS TO CPN60 IN THE PRESENCE OF MG-ATP AND SUPPRESS
C1 THE ATPASE ACTIVITY OF THE LATTER (BY SIMILARITY).
C2 SUBUNIT: REPETITION OF 7 SUBUNITS ARRANGED IN A RING (BY
C3 SIMILARITY).
C4 SIMILARITY: BELONGS TO THE GROES CHAPERONIN FAMILY.
C5 EMBL: AF010251; AAB56325.1;
C6 PROSITE: P500691; CHAPERONINS_CPNI0; 1.
C7 PFAM: PF00165; CPN10; 1.
C8 PRINTS: PR02977; CHAPERONIN10.
C9 Chaperone.
C0 SEQUENCE 93 AA: 10040 MW: 35954571 CRC32:

Query Match 58.8% Score 67; DB 2; Length 93;
Best Local Similarity 35.7% Pred. No. 4,81e-01;
Matches 5; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
DB 79 CDYLVSARDVLAV 14
Q: 1 EYLILSARDVLAV 14

RESULT 7
ID Q67942 PRELIMINARY: PRT: 122 AA.
AC Q67942
DI 01-NOV-1999 (TRENBLrel. 07, Created)
DI 01-NOV-1999 (TRENBLrel. 07, Last sequence update)
DI 01-NOV-1999 (TRENBLrel. 12, Last annotation update)
GN MOPE.
OS Aquifex aeolicus.
OT Bacteria; Aquificales; Aquificaceae; Aquifex
RN [1]
RP SEQUENCE FROM N.A.
R1 STRAIN-VF5.
R2 MEDLINE: 98155666.
R3 DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
R4 GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,
R5 FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.,
R6 Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
R7 The complete genome of the hyperthermophilic bacterium Aquifex
R8 aeolicus.
R9 Nature 392:353-358(1998).
RN [2]
RP SEQUENCE FROM N.A.
R1 STRAIN-VF5.
R2 DECKERT G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
R3 GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AUJAY M., HUBER R.,
R4 FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.,
R5 Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
R6 EMBL: AF000777; AAC07898.1;
R7 PFAM: PF00165; CPN10;
R8 PRINTS: PR02977; CHAPERONIN10.
R9 Chaperone.
R0 SEQUENCE 122 AA: 13639 MW: 44277AAB CRC32:
Query Match 58.8% Score 67; DB 2; Length 122;

Best Local Similarity 61.5% Pred. No. 4,81e-01;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

DB 83 YLWSEDEVAVV 95
Q: 3 YLILSARDVLAV 15
RESULT 8
ID Q9W634 PRELIMINARY: PRT: 304 AA.
AC Q9W634
DI 01-NOV-1999 (TRENBLrel. 12, Created)
DI 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DI 01-NOV-1999 (TRENBLrel. 12, Last annotation update)
DE MRNA, PARTIAL CDS, SIMILAR TO MOUSE IMPRINTED
DE GENE IMPACT (FRAGMENT).
DE Xenopus borealis (Kenyan clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae; Xenopodinae;
OC Xenopus.
RN [1]
RP SEQUENCE FROM N.A.
R1 YAMADA Y.;
R2 "Xenopus borealis homologue of mouse imprinted gene impact".
R3 Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
R4 EMBL: AB021258; BAA78421.1;
R5 NON-TER;
R6 NON-TER;
R7 NON-TER;
R8 SEQUENCE 304 AA: 34661 MW: D94F43B1 CRC32:

Query Match 58.8% Score 67; DB 13; Length 304;
Best Local Similarity 69.2% Pred. No. 4,81e-01;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

DB 248 ILDARDVMVVS 260
Q: 5 ILSARDVLAVSK 17
RESULT 9
ID Q9W625 PRELIMINARY: PRT: 312 AA.
AC Q9W625
DI 01-NOV-1999 (TRENBLrel. 12, Created)
DI 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DI 01-NOV-1999 (TRENBLrel. 12, Last annotation update)
DE XIMPACT.
DE Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia;
OC Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae; Xenopodinae;
OC Xenopus.
RN [1]
RP SEQUENCE FROM N.A.
R1 MEDLINE: 99167356.
R2 YAMADA Y., HAGWARA Y., SHICKAWA K., SAKAKI Y., ITO T.,
R3 "spatiotemporal, allelic, and enforced expression of Ximpact, the
R4 Xenopus homolog of mouse imprinted gene impact".
R5 Biochem. Biophys. Res. Commun. 258:162-169(1999).
R6 EMBL: AB020319; BAA78409.1;
R7 SEQUENCE 312 AA: 35640 MW: QC25D754 CRC32:

Query Match 58.8% Score 67; DB 13; Length 312;
Best Local Similarity 69.2% Pred. No. 4,81e-01;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

DB 251 ILDARDVMVVS 263
Q: 5 ILSARDVLAVSK 17
RESULT 10
ID Q94889 PRELIMINARY: PRT: 465 AA.
AC Q94889
DI 01-MAY-1999 (TRENBLrel. 10, Created)

```

DT 01-MAY-1999 (TREMBLER: 10, Last sequence update)
DI 01-MAY-1999 (TREMBLER: 10, Last annotation update)
DE KIAA0795 PROTEIN (FRAGMENT)
GN KIAA0795
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Euteria; Primates; Catarrhini; Homidae; Homo.
RN [1]
RT SEQUENCE FROM N.A.
RC TISSUE-BRAIN
RA MEDLINE: 9907487.
RA NAGASE T., ISHIKAWA K., SUYAMA M., KIKUNO R., MIYAJIMA N., TANAKA A.,
RA KOTANI H., NOMURA N., GHARA O.
RA *Prediction of the coding sequences of unidentified human genes. XI.
RI The complete sequences of 100 new cDNA clones from brain which code
RI for large proteins in vitro.
RL CNA Res. 5:277-286(1998)
DR EMBL: AB018319; BAA34315.1; -.
FT NON-TER
SQ SEQUENCE 465 AA: 51392 MW: 55159266 CRC32:

Query Match 57.0% Score 65: DB 4: Length 465:
Best Local Similarity 60.0% Pred. No. 1,140-00:
Matches 9: Conservative 4: Mismatches 4: Indels 0: Gaps 0:

DB 65 EBYLMSRDLVLSR 86
QY 1 EBYLMSRDLVLSR 17

RESULT 1:
ID 066201 PRELIMINARY: PRT: 88 AA
AC 066201
DI 01-AUG-1998 (TREMBLER: 07, Created)
DI 01-AUG-1998 (TREMBLER: 07, Last sequence update)
DE 01-MAY-1999 (TREMBLER: 10, Last annotation update)
DE SIMILAR TO GROS PROTEIN (FRAGMENT)
OS Seriatia rubidaea.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Seriatia.
RN [1]
RT SEQUENCE FROM N.A.
RC STRAIN-JCM 1240:
RA HARADA H., ISHIKAWA H.
RA J. Gen. Appl. Microbiol. 43:355-361(1997).
DR EMBL: AB008143; BAA25218.1; -.
DR PFAM: PF00166; cpn10.1; -.
FT NON-TER
SQ SEQUENCE 84 AA: 9212 MW: 25199566 CRC32:

Query Match 57.0% Score 65: DB 2: Length 86:
Best Local Similarity 60.0% Pred. No. 1,140-00:
Matches 9: Conservative 3: Mismatches 3: Indels 0: Gaps 0:

DB 72 EBYLMSRDLVLSR 86
QY 1 EBYLMSRDLVLSR 15

RESULT 12
ID 066189 PRELIMINARY: PRT: 88 AA
AC 066189
DI 01-AUG-1998 (TREMBLER: 07, Created)
DI 01-AUG-1998 (TREMBLER: 07, Last sequence update)
DE 01-MAY-1999 (TREMBLER: 10, Last annotation update)
DE SIMILAR TO GROS PROTEIN (FRAGMENT)
OS Enterobacter asburiae.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Enterobacter.
RN [1]
RT SEQUENCE FROM N.A.
RC STRAIN-JCM6051:
RA HARADA H., ISHIKAWA H.

```

```

RL J. Gen. Appl. Microbiol. 43:355-361(1997).
DR EMBL: AB008137; BAA25206.1; -.
DR PFAM: PF00166; cpn10.1; -.
FT NON-TER
SQ SEQUENCE 88 AA: 9254 MW: 25045453 CRC32:

Query Match 57.0% Score 65: DB 2: Length 86:
Best Local Similarity 60.0% Pred. No. 1,140-00:
Matches 9: Conservative 4: Mismatches 3: Indels 0:

DB 72 EBYLMSRDLVLSR 86
QY 1 EBYLMSRDLVLSR 15

RESULT 13
ID 066211 PRELIMINARY: PRT: 88 AA
AC 066211
DI 01-AUG-1998 (TREMBLER: 07, Created)
DI 01-AUG-1998 (TREMBLER: 07, Last sequence update)
DE 01-MAY-1999 (TREMBLER: 10, Last annotation update)
DE SIMILAR TO GROS PROTEIN (FRAGMENT)
OS Klebsiella planticola
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Klebsiella.
RN [1]
RT SEQUENCE FROM N.A.
RC STRAIN-JCM 7251:
RA HARADA H., ISHIKAWA H.
RA J. Gen. Appl. Microbiol. 43:355-361(1997).
DR EMBL: AB008148; BAA25228.1; -.
DR PFAM: PF00166; cpn10.1; -.
FT NON-TER
SQ SEQUENCE 88 AA: 9212 MW: 25199566 CRC32:

Query Match 57.0% Score 65: DB 2: Length 86:
Best Local Similarity 60.0% Pred. No. 1,140-00:
Matches 9: Conservative 3: Mismatches 3: Indels 0: Gaps 0:

DB 72 EBYLMSRDLVLSR 86
QY 1 EBYLMSRDLVLSR 15

RESULT 14
ID 066193 PRELIMINARY: PRT: 88 AA
AC 066193
DI 01-AUG-1998 (TREMBLER: 07, Created)
DI 01-AUG-1998 (TREMBLER: 07, Last sequence update)
DE 01-MAY-1999 (TREMBLER: 10, Last annotation update)
DE SIMILAR TO GROS PROTEIN (FRAGMENT)
OS Enterobacter quercusiae
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Enterobacter.
RN [1]
RT SEQUENCE FROM N.A.
RC STRAIN-JCM 1234:
RA HARADA H., ISHIKAWA H.
RA J. Gen. Appl. Microbiol. 43:355-361(1997).
DR EMBL: AB008139; BAA25210.1; -.
DR PFAM: PF00166; cpn10.1; -.
FT NON-TER
SQ SEQUENCE 88 AA: 9239 MW: 25974000 CRC32:

Query Match 57.0% Score 65: DB 2: Length 86:
Best Local Similarity 60.0% Pred. No. 1,140-00:
Matches 9: Conservative 3: Mismatches 3: Indels 0: Gaps 0:

DB 72 EBYLMSRDLVLSR 86
QY 1 EBYLMSRDLVLSR 15

```

```


RESULT 15
ID O66219 PRELIMINARY: PRT: 88 AA
AC O66219
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DI 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
CI 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE SIMILAR TO GROS PROTEIN (FRAGMENT).
CS Erwinia carotovora.
OC Bacteria: Proteobacteria: gamma subdivision: Enterobacteriaceae:
OC Erwinia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAINS:AM 12433.
RA HARADA H., ISHIKAWA H.:
RC J. Gen. Appl. Microbiol. 43:355-361(1997).
DR EMBL: AB008152; BAA25236.1;
DR PFAM: PF00166; cpn10; 1.
FT NON-TER 1
SQ SEQUENCE 88 AA: 9233 MW: 0ADDE5CE CRC32:

Query Match 57.0% Score 65: DB 2: length 88:
Best Local Similarity 60.0% Pred. NC. 1.14e-00:
Matches 9: Conservative 3: Mismatches 3: Indels 0: Gaps 0:

DA 72 REVLIMSESDILAV 86
II I I I I I I I I
QY 1 REVLISARDVAVV 15

Search completed: Mon Jun 19 16:07:23 2000
Job time : 12 secs.

```



Release 3.1A John F. Collins, BioComputing Research Unit,
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Wsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Jun 19 16:05:10 2000: MasPar time 2.59 Seconds

Tabular output not generated. 82.425 Million cell updates/sec

Title: >US-09-142-524A-7

Description: (1-9) from US09142524A.pep

Sequence: 1 FIKRVSNVI 9

Scoring table: FAM 150
Gap 15

Searched: 185953 seqs, 23685106 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries
Maximum DB seq length 9

Database: a:geneseq35
l:geneseqf

Statistics: Mean 16.765, Variance 46.25 scale 0.362

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB ID | Description | Pred. No. |
|------------|-------|-------|--------|-------|-------------------------------|-----------|
| 1 | 54 | 100.0 | 9 | 1 | W57751 Residues 109-117 of Cr | 1.25e-00 |

Note: Post-processor removed 44 summaries from list due to search parameters chosen.

ALIGNMENTS

RESULT 1
 ID W57751 standard: peptide 9 AA.
 AC W57751
 DI 17-SEP-1998 (first entry)
 DE Residues 109-117 of Cr
 KW Cr 109-117 of Cr
 KW HLA class II molecule.
 OS Cryptomeria japonica.
 PN W09620902.AV
 PC 22-MAY-1998
 FR 12-NOV-1997 J04129
 PR 13-NOV-1996 JP-102653.
 PA (MEIP) MEIJ: MILK PRO CO LTD.
 PI Saitaki K, Kuno K, Kume A, Sone T.
 DR WPI: 96-24761/25.

PT Peptides derived from Japanese cedar pollen antigens are
 PT immunotherapeutic agents - useful for allergy treatment and typing
 PT HLA class II molecules in allergy sufferers
 PS Claim 12: Page 26: 50pp: Japanese
 CC This sequence represents residues 109-117 of the Cr 109-117 protein and
 CC is a peptide of the invention. The peptides are derived from Japanese
 CC cedar pollen antigens, and are used as immunotherapeutic agents in the
 CC treatment of allergy. The peptides can be used for identification and
 CC typing of the particular HLA class II molecules in an allergy sufferer,
 CC and also for peptide immunotherapy of an allergy. Using these peptides
 CC the immunotherapy can be targeted more specifically to the requirements
 CC of the individual patient, allowing more effective treatment of an
 CC allergy, including those patients for whom treatment with a conventional
 CC immunotherapeutic agent is ineffective.
 SQ Sequence 9 AA:

Query Match 100.0%: Score 54: DB 1: Length 9:

Best Local Similarity 100.0%: Pred. No. 1.25e-00:
 Matches 9: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

DB 1 FIKRVSNVI 9
 QY 1 FIKRVSNVI 9

Search completed: Mon Jun 19 16:05:22 2000

Job time : 12 secs.

WATERBURY (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit,
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution Rights by Oxford Molecular Ltd

Mysearch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Tue Jun 20 13:39:14 2000; Waspar time 2.19 Seconds
59.362 Million cell updates/sec

Tubular output not generated.

Title: >US-09-142-524A-7
Description: (1-9) from US09142524A.pep
Perfect Score: 64
Sequence: 1 FIKRVSNVI 9

Scoring table: PAM 150
Gap 15

Searched: 145341 seqs, 14437480 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries
Maximum DB seq length 9

Database: a-issued
1:5A_COMB 2:5B_COMB 3:6_COMB 4:PCT_COMB 5:backfiles1

Statistics: Mean 15.775; Variance 46.625; scale 0.338

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result | Query | | | | |
|--------|-------|--------------|-------|-------------|----------|
| No. | Score | Match length | DB ID | Description | Pred No. |

No matches found.

Search completed: Tue Jun 20 13:39:17 2000
Job time : 3 secs.

.....
[W][I][T][E][R][M][A][N] (TM)
.....

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Wierch_pp prote.n - protein database search, using Smith-Waterman algorithm

Ran on: Mon Jun 19 16:08:54 2000: Maspar time 13.59 Seconds

Linear output not generated. 66.717 Million cell updates/sec

Title: >US-09-142-524A-7
Description: (1-9) from US09142524A.pdf
Perfect Score: 64
Sequence: 1 FIRRVSNVI 9

Scoring table: PAM 150
Gap 15

Searched: 721208 seqs, 100765575 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries
Maximum DB seq length 9

Database: a:pending
1:FT 2:06 3:060 4:07 5:080 6:081 7:082 8:083 9:084A
10:084B 11:085 12:086 13:08 14:088 15:089 16:090 17:091
18:092 19:093 20:094 21:095 22:NEWP 23:NEW060 24:NEW08
25:NEW09

Statistics: Mean 18.575: Variance 43.933: scale 0.423

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result | Query | No. | Score | Match | Length | DB | ID | Description | Pred. No. |
|--------|-------|-----|-------|-------|--------|----|----|-------------|-----------|
|--------|-------|-----|-------|-------|--------|----|----|-------------|-----------|

No matches found.

Search completed: Mon Jun 19 16:09:17 2000
J.E. time : 23 secs.

.....
MPSRELA (TM)
.....

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd
MPSrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Mon Jun 19 16:04:38 2000 MasPar time 4.43 Seconds
95.841 Million cell updates/sec

Tabular output not generated.

Title: >US-09-142-524A-7
Description: (179) from USC9142524A.pep
Perfect Score: 54
Sequence: 1 FKKVSNV: 9

Scoring table: PAM 150
Gap 15

Searched: 142080 seqs, 47172406 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries
Maximum DB seq length 9

Database: pir62
1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 23.798; Variance 28.919; scale 0.823

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result | Query | | | | |
|--------|-------|-------|--------|-------|-------------|
| No. | Score | Match | Length | DB ID | Description |
| | | | | | Pred. No. |

No matches found.

Search completed: Mon Jun 19 16:04:53 2000
Job time: 15 secs.

[WATERMAN] (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsearch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Jun 19 16:03:40 2000: MasPar time 3.37 Seconds
81.234 Million cell updates/sec

Tabular output not generated.

Title: >US-09-142-524A-7
Description: (1-9) from USC9142524A.pep
Perfect score: 64
Sequence: 1 FIKRVSNVI 9

Scoring table: PAM 150
Gap 15

Searched: 83857 seqs, 30454973 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries
Maximum DB seq length 9

Database: swiss-prot38
1:swissprot

Statistics: Mean 24.343: Variance 25.165: scale 0.967

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result | Query | | | | |
|--------|-------|-------|--------|-------|-------------|
| No. | Score | Match | Length | DB ID | Description |
| | | | | | Prod. No. |

No matches found.

Search completed: Mon Jun 19 16:03:47 2000
Job time : 7 secs.

(M2)

... Ltd

mpsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Jun 19 16:04:03 2000; MasPar time 8.30 Seconds

Tabular output not generated.

Title: >UJS-GG-142-524A-7

Description: (1-9) from US09142524A pep

perfect score:

Sequence: 1 F:KRVSNV: 9

Scoring table: PAM 150

Gap 15

Searched: 225878 seqs, 69334122 residues

Post-processing: Minimum Match 08

Listing first 45 summaries

Maximum DB seq length: 9

Database: sptrembl12

1:sp_arctea 2:sp_bacteria 3:sp_fungi 4:sp_human
5:sp_invertebrate 6:sp_mammal 7:sp_mhc 8:sp_organelle
9:sp_phase 10:sp_plant 11:sp_rodent 12:sp_unclassified
13:sp_vertebrate 14:sp_virus

| | |
|-------------|---|
| Statist.cs. | Mean 23.348; Variance 25.366; scale 0.920 |
|-------------|---|

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being paired, and is derived by analysis of the total score distribution.

SUMMARY

| Result No. | Score | Query Match | DB ID | Description | Pred. No. |
|------------|-------|-------------|-------|-------------|-----------|
| 1 | 100 | 100 | 1 | 1 | 1 |
| 2 | 100 | 100 | 2 | 2 | 2 |
| 3 | 100 | 100 | 3 | 3 | 3 |
| 4 | 100 | 100 | 4 | 4 | 4 |
| 5 | 100 | 100 | 5 | 5 | 5 |
| 6 | 100 | 100 | 6 | 6 | 6 |
| 7 | 100 | 100 | 7 | 7 | 7 |
| 8 | 100 | 100 | 8 | 8 | 8 |
| 9 | 100 | 100 | 9 | 9 | 9 |
| 10 | 100 | 100 | 10 | 10 | 10 |
| 11 | 100 | 100 | 11 | 11 | 11 |
| 12 | 100 | 100 | 12 | 12 | 12 |
| 13 | 100 | 100 | 13 | 13 | 13 |
| 14 | 100 | 100 | 14 | 14 | 14 |
| 15 | 100 | 100 | 15 | 15 | 15 |
| 16 | 100 | 100 | 16 | 16 | 16 |
| 17 | 100 | 100 | 17 | 17 | 17 |
| 18 | 100 | 100 | 18 | 18 | 18 |
| 19 | 100 | 100 | 19 | 19 | 19 |
| 20 | 100 | 100 | 20 | 20 | 20 |
| 21 | 100 | 100 | 21 | 21 | 21 |
| 22 | 100 | 100 | 22 | 22 | 22 |
| 23 | 100 | 100 | 23 | 23 | 23 |
| 24 | 100 | 100 | 24 | 24 | 24 |
| 25 | 100 | 100 | 25 | 25 | 25 |
| 26 | 100 | 100 | 26 | 26 | 26 |
| 27 | 100 | 100 | 27 | 27 | 27 |
| 28 | 100 | 100 | 28 | 28 | 28 |
| 29 | 100 | 100 | 29 | 29 | 29 |
| 30 | 100 | 100 | 30 | 30 | 30 |
| 31 | 100 | 100 | 31 | 31 | 31 |
| 32 | 100 | 100 | 32 | 32 | 32 |
| 33 | 100 | 100 | 33 | 33 | 33 |
| 34 | 100 | 100 | 34 | 34 | 34 |
| 35 | 100 | 100 | 35 | 35 | 35 |
| 36 | 100 | 100 | 36 | 36 | 36 |
| 37 | 100 | 100 | 37 | 37 | 37 |
| 38 | 100 | 100 | 38 | 38 | 38 |
| 39 | 100 | 100 | 39 | 39 | 39 |
| 40 | 100 | 100 | 40 | 40 | 40 |
| 41 | 100 | 100 | 41 | 41 | 41 |
| 42 | 100 | 100 | 42 | 42 | 42 |
| 43 | 100 | 100 | 43 | 43 | 43 |
| 44 | 100 | 100 | 44 | 44 | 44 |
| 45 | 100 | 100 | 45 | 45 | 45 |
| 46 | 100 | 100 | 46 | 46 | 46 |
| 47 | 100 | 100 | 47 | 47 | 47 |
| 48 | 100 | 100 | 48 | 48 | 48 |
| 49 | 100 | 100 | 49 | 49 | 49 |
| 50 | 100 | 100 | 50 | 50 | 50 |
| 51 | 100 | 100 | 51 | 51 | 51 |
| 52 | 100 | 100 | 52 | 52 | 52 |
| 53 | 100 | 100 | 53 | 53 | 53 |
| 54 | 100 | 100 | 54 | 54 | 54 |
| 55 | 100 | 100 | 55 | 55 | 55 |
| 56 | 100 | 100 | 56 | 56 | 56 |
| 57 | 100 | 100 | 57 | 57 | 57 |
| 58 | 100 | 100 | 58 | 58 | 58 |
| 59 | 100 | 100 | 59 | 59 | 59 |
| 60 | 100 | 100 | 60 | 60 | 60 |
| 61 | 100 | 100 | 61 | 61 | 61 |
| 62 | 100 | 100 | 62 | 62 | 62 |
| 63 | 100 | 100 | 63 | 63 | 63 |
| 64 | 100 | 100 | 64 | 64 | 64 |
| 65 | 100 | 100 | 65 | 65 | 65 |
| 66 | 100 | 100 | 66 | 66 | 66 |
| 67 | 100 | 100 | 67 | 67 | 67 |
| 68 | 100 | 100 | 68 | 68 | 68 |
| 69 | 100 | 100 | 69 | 69 | 69 |
| 70 | 100 | 100 | 70 | 70 | 70 |
| 71 | 100 | 100 | 71 | 71 | 71 |
| 72 | 100 | | | | |

No matches found.

Search completed: Mon Jun 19 16:04:22 2000

SEARCH COMPLETED: 19
Job time : 19 secs.

PT Peptides derived from Japanese cedar pollen antigens are immunotherapeutic agents - useful for allergy treatment and typing
FI HLA class II molecules in allergy sufferers
PS Claim 12: Page 28: 50pp: Japanese
CC This sequence represents residues 211-225 of the Cry j 1 protein, and is a peptide of the invention. The peptides are derived from Japanese cedar pollen antigens, and are used as immunotherapeutic agents in the treatment of allergy. The peptides can be used for identification and typing of the particular HLA class II molecules in an allergy sufferer, and also for peptide immunotherapy of an allergy. Using these peptides, the immunotherapy can be targeted more specifically to the requirements of the individual patient, allowing more effective treatment of an allergy, including those patients for whom treatment with a conventional immunotherapeutic agent is ineffective.
SQ Sequence 15 AA:

Query Match 100.0% Score 106 DB 1 Length 15
Best Local Similarity 100.0% Pred. No. 8.38e-05
Matches 15: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

DB 1 KSMKVIVAFNQFGPN 15
QY 1 KSMKVIVAFNQFGPN 15

RESULT 3
ID W80349 standard: peptide: 17 AA.

AC W80349
DE 11-JAN-1999 (first entry)
KW Sugi allergen protein Cryj2 derived epitope for T cells.
KW T cell epitope: sugi allergen proteins Cryj1; Cryj2; treatment: sugi-pollinosis; allergic reaction: pollen.
OS Synthetic.
PN J10259198-A
PR 29-SEP-1998
PF 22-DEC-1997: 353448.
PA (HAYB) HAYASHIBARA SEIBUTSU KAKJAKU.
PA (SANY) SANKYO CO LTD.
DE WPI: 94-57707/49.
PT A linked T cell epitope peptide - used for the treatment of sugi-pollinosis.
PS Claim 7: Page 18: 21pp: Japanese.
CC W80339-59 represent epitopes for T cells, derived from the sugi allergen proteins Cryj1 (W80339-44), W80350-53 and W80345-49 and W80354-61. The peptides are useful for the treatment of sugi-pollinosis, an allergic reaction of the body to pollen.
SQ Sequence 17 AA:

Query Match 100.0% Score 106 DB 1 Length 17
Best Local Similarity 100.0% Pred. No. 8.38e-05
Matches 15: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

DB 2 KSMKVIVAFNQFGPN 16
QY 1 KSMKVIVAFNQFGPN 15

RESULT 4
ID W42144 standard: peptide: 20 AA.

AC W42144
DE 16-JUN-1998 (first entry)
KW T-cell epitope peptide 24 from Japanese cypress pollen antigen Chaol.
KW Japanese cypress pollen; antigen: T-cell epitope: Chaol; Chaol2; diagnosis: allergy; spring tree pollen disease; pollinosis.
OS Chamaecyparis obtusa.
PN W09747648-A1.
PR 18-DEC-1997.
PF 12-JUN-1997: 502031.
PA (MELP) MELP MILK PROD CO LTD.
PA Dairiki K. Kiko K.
DE WPI: 96-052242/05.

PT T-cell epitope peptide portion of Japanese cypress pollen antigens Chaol and Chaol2 - used for diagnosis and treatment of spring tree pollen disease
PS Claim 1: Page 27-28: 71pp: Japanese
CC The present sequence represents a T-cell epitope peptide from Japanese cypress pollen antigen Chaol. The present invention describes peptides which correspond to the T-cell epitope sites on Japanese cypress pollen antigens Chaol and Chaol2. The peptides can be used as a reagent for the diagnosis of allergy to Japanese cypress pollen, and as an antigen in the treatment and prevention of spring tree pollen disease in which the pollinosis involves reactivity to Japanese cypress pollen.
SQ Sequence 20 AA:

Query Match 100.0% Score 106 DB 1 Length 20
Best Local Similarity 100.0% Pred. No. 8.38e-05
Matches 15: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

DB 1 KSMKVIVAFNQFGPN 15
QY 1 KSMKVIVAFNQFGPN 15

RESULT 5
ID R82512 standard: Protein: 20 AA.

AC R82512:
DE 15-APR-1996 (first entry)
KW Cry j 1 Japanese Cedar pollen allergen peptide fragment (801-22).
KW Cry j 1 Japanese cedar pollen allergen, modified, drug production.
KW allergy: Cryptomeria japonica.
CS Cryptomeria japonica.
PN W09527796-A.
PD 19-OCT-1995.
PF 06-APR-1995: 004245.
PR 08-APR-1994: US-226248.
PR 06-DEC-1994: US-350235.
PA (IMMU) IMMULOGIC PHARM CORP.
PI Chen X; Evans S; Franzen HM; Kuo M; Powers SP; Shaked Z.
DE WPI: 95-366391/47.
PT Modified Cryptomeria japonica (Cry j) 1 peptide(s) - useful for treating allergy to Japanese cedar pollen allergen or immunologically cross reactive allergens.
PS Disclosure: Figure 2: 60pp: English.
CC Novel peptides of cry j 1 have been modified as a part of a preformulation scheme to develop an optimised drug product for therapeutic treatment of humans suffering from allergy to Japanese cedar pollen allergen or an allergen which is immunologically cross reactive with Japanese cedar pollen allergen. Such modified peptides possess certain characteristics which render them particularly suitable for drug product formulation. Peptide fragments of Cry j 1, modified and unmodified, are given in W8491-R82525. This peptide fragment corresponds to amino acids 211-230 of the allergen mature protein.
CC protein.
SQ Sequence 20 AA:

Query Match 100.0% Score 106 DB 1 Length 20
Best Local Similarity 100.0% Pred. No. 8.38e-05
Matches 15: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

DB 1 KSMKVIVAFNQFGPN 15
QY 1 KSMKVIVAFNQFGPN 15

RESULT 6
ID R45563 standard: Protein: 20 AA.

AC R45563:
DE 13-JUL-1994 (first entry)
KW Cry j 1 pollen allergen peptide CUI-22.
KW Japanese cedar; detection: allergy; treatment: diagnosis: T cell epitope; sensitivity.
OS Cryptomeria japonica.
PN W09401560-A.

PD 20-JAN-1994.
 PF 15-JAN-1993; U00139.
 PR 10-JUL-1992; WO-005661.
 PA 01-SEP-1992; US-938990.
 PA (IMMU-) IMMUNOLOGIC PHARM CORP.
 PI Bond JF, Garman RD, Griffith IJ, Kuo M, Pollock J.
 DR WPI: 94-035066/04.
 PI Antigens derived from Japanese cedar pollen allergen Cry j 1.
 PI contain at least two T cell epitope(s), used to treat or diagnose
 PI allergy.
 PS Claim 76: Fig 13: 137pp: English.
 CC The sequence is that of an isolated peptide of the Japanese cedar
 CC pollen allergen Cry j 1 (amino acids 211-230). The peptide, CJI-43.8,
 CC can be used for the treatment and diagnosis of allergies associated
 CC with Japanese cedar pollen. It has enhanced therapeutic properties
 CC but reduced side effects compared to naturally occurring allergens.
 SQ Sequence 26 AA:

Query Match 100.0%; Score 106; DB 1; Length 20;
 Best Local Similarity 100.0%; Pred. No. 8.38e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 KSMKVTVAEFGGPN 15
 QY 1 KSMKVTVAEFGGPN 15

RESULT 7
 ID R45592 standard; Protein: 26 AA.
 AC R45592.
 DT 13-JUL-1994 (first entry)
 DE Cry j 1 pollen allergen peptide CJI-43.8.
 KW Japanese cedar; detection: allergy; treatment: diagnosis;
 KW T cell epitope; sensitivity.
 OS Cryptomeria japonica.
 PN WO9401560-A.
 PD 20-JAN-1994.
 PF 15-JAN-1993; U00139.
 PR 10-JUL-1992; WO-005661.
 PA 01-SEP-1992; US-938990.
 PA (IMMU-) IMMUNOLOGIC PHARM CORP.
 PI Bond JF, Garman RD, Griffith IJ, Kuo M, Pollock J.
 DR WPI: 94-035066/04.
 PI Antigens derived from Japanese cedar pollen allergen Cry j 1.
 PI contain at least two T cell epitope(s), used to treat or diagnose
 PI allergy.
 PS Claim 76: Fig 13: 137pp: English.
 CC The sequence is that of an isolated peptide of the Japanese cedar
 CC pollen allergen Cry j 1. The peptide, CJI-43.8, can be used for
 CC the treatment and diagnosis of allergies associated with Japanese
 CC cedar pollen. It has enhanced therapeutic properties but reduced
 CC side effects compared to naturally occurring allergens.
 SQ Sequence 26 AA:

Query Match 100.0%; Score 106; DB 1; Length 26;
 Best Local Similarity 100.0%; Pred. No. 8.38e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 KSMKVTVAEFGGPN 15
 QY 1 KSMKVTVAEFGGPN 15

RESULT 8
 ID R45591 standard; Protein: 26 AA.
 AC R45591.
 DT 13-JUL-1994 (first entry)
 DE Cry j 1 pollen allergen peptide CJI-43.9.
 KW Japanese cedar; detection: allergy; treatment: diagnosis;
 KW T cell epitope; sensitivity.
 OS Cryptomeria japonica.
 PN WO9401560-A.
 PD 20-JAN-1994.

PF 15-JAN-1993; U00139.
 PR 10-JUL-1992; WO-005661.
 PA 01-SEP-1992; US-938990.
 PA (IMMU-) IMMUNOLOGIC PHARM CORP.
 PI Bond JF, Garman RD, Griffith IJ, Kuo M, Pollock J.
 DR WPI: 94-035066/04.
 PI Antigens derived from Japanese cedar pollen allergen Cry j 1.
 PI contain at least two T cell epitope(s), used to treat or diagnose
 PI allergy.
 PS Claim 76: Fig 13: 137pp: English.
 CC The sequence is that of an isolated peptide of the Japanese cedar
 CC pollen allergen Cry j 1. The peptide, CJI-43.9, can be used for
 CC the treatment and diagnosis of allergies associated with Japanese
 CC cedar pollen. It has enhanced therapeutic properties but reduced
 CC side effects compared to naturally occurring allergens.
 SQ Sequence 26 AA:

Query Match 100.0%; Score 106; DB 1; Length 26;
 Best Local Similarity 100.0%; Pred. No. 8.38e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 KSMKVTVAEFGGPN 15
 QY 1 KSMKVTVAEFGGPN 15

RESULT 9
 ID R45594 standard; Protein: 26 AA.
 AC R45594.
 DT 13-JUL-1994 (first entry)
 DE Cry j 1 pollen allergen peptide CJI-43.12.
 KW Japanese cedar; detection: allergy; treatment: diagnosis;
 KW T cell epitope; sensitivity.
 OS Cryptomeria japonica.
 PN WO9401560-A.
 PD 20-JAN-1994.
 PF 15-JAN-1993; U00139.
 PR 10-JUL-1992; WO-005661.
 PA 01-SEP-1992; US-938990.
 PA (IMMU-) IMMUNOLOGIC PHARM CORP.
 PI Bond JF, Garman RD, Griffith IJ, Kuo M, Pollock J.
 DR WPI: 94-035066/04.
 PI Antigens derived from Japanese cedar pollen allergen Cry j 1.
 PI contain at least two T cell epitope(s), used to treat or diagnose
 PI allergy.
 PS Claim 76: Fig 13: 137pp: English.
 CC The sequence is that of an isolated peptide of the Japanese cedar
 CC pollen allergen Cry j 1. The peptide, CJI-43.12, can be used for
 CC the treatment and diagnosis of allergies associated with Japanese
 CC cedar pollen. It has enhanced therapeutic properties but reduced
 CC side effects compared to naturally occurring allergens.
 SQ Sequence 26 AA:

Query Match 100.0%; Score 106; DB 1; Length 26;
 Best Local Similarity 100.0%; Pred. No. 8.38e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 KSMKVTVAEFGGPN 15
 QY 1 KSMKVTVAEFGGPN 15

RESULT 10
 ID R45590 standard; Protein: 28 AA.
 AC R45590.
 DT 13-JUL-1994 (first entry)
 DE Cry j 1 pollen allergen peptide CJI-43.8.
 KW Japanese cedar; detection: allergy; treatment: diagnosis;
 KW T cell epitope; sensitivity.
 OS Cryptomeria japonica.
 PN WO9401560-A.
 PD 20-JAN-1994.
 PF 15-JAN-1993; U00139.

10-JUL-1992: WO-005661.
 PR 01-SEP-1992: US-938990.
 PA (IMMU-) IMMUNOLOGIC PHARM CORP.
 PI Bond JF, Garman RD, Griffith LJ, Kuo M, Pollock J;
 DR WPI: 94-035066/04.
 DE Antigen derived from Japanese cedar pollen allergen Cry j 1.
 PI Antigen derived from Japanese cedar pollen allergen Cry j 1.
 PI contain at least two T cell epitope(s), used to treat or diagnose
 PT allergy.
 PS Claim 76: Fig 18: 137pp: English.
 CC The sequence is that of an isolated peptide of the Japanese cedar
 CC pollen allergen Cry j 1. The peptide, CJI-43.8, can be used for
 CC the treatment and diagnosis of allergies associated with Japanese
 CC cedar pollen. It has enhanced therapeutic properties but reduced
 CC side effects compared to naturally occurring allergens.
 SQ Sequence 28 AA.

Query Match 100.0% Score 106; DB 1; Length 28;
 Best Local Similarity 100.0% Pred. No. 8.38e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DE 1 KSMKVTVAFNQFGPN 15
 QY 1 KSMKVTVAFNQFGPN 15

RESULT 11
 ID R45593 standard: Protein: 28 AA.

AC R45593:
 DT 13-JUL-1994 (first entry)
 DE Cry j 1 pollen allergen peptide CJI-43.11.
 KW Japanese cedar; detection: allergy; treatment: diagnosis;
 KW T cell epitope; sensitivity.
 OS Cryptomeria japonica.
 PN WO9401560-A.
 PD 20-JAN-1994.
 PF 15-JAN-1993; U00139.
 PR 10-JUL-1992: WO-005661.
 PA (IMMU-) IMMUNOLOGIC PHARM CORP.
 PI Bond JF, Garman RD, Griffith LJ, Kuo M, Pollock J;
 DR WPI: 94-035066/04.
 PI Antigen derived from Japanese cedar pollen allergen Cry j 1.
 PI contain at least two T cell epitope(s), used to treat or diagnose
 PT allergy.
 PS Claim 76: Fig 18: 137pp: English.
 CC The sequence is that of an isolated peptide of the Japanese cedar
 CC pollen allergen Cry j 1. The peptide, CJI-43.11, can be used for
 CC the treatment and diagnosis of allergies associated with Japanese
 CC cedar pollen. It has enhanced therapeutic properties but reduced
 CC side effects compared to naturally occurring allergens.
 SQ Sequence 28 AA.

Query Match 100.0% Score 106; DB 1; Length 28;
 Best Local Similarity 100.0% Pred. No. 8.38e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DE 1 KSMKVTVAFNQFGPN 15
 QY 1 KSMKVTVAFNQFGPN 15

RESULT 12
 ID R45589 standard: Protein: 28 AA.

AC R45589:
 DT 13-JUL-1994 (first entry)
 DE Cry j 1 pollen allergen peptide CJI-43.7.
 KW Japanese cedar; detection: allergy; treatment: diagnosis;
 KW T cell epitope; sensitivity.
 OS Cryptomeria japonica.
 PN WO9401560-A.
 PD 20-JAN-1994.
 PF 15-JAN-1993; U00139.
 PR 10-JUL-1992: WO-005661.

01-SEP-1992: US-938990.
 PR (IMMU-) IMMUNOLOGIC PHARM CORP.
 PI Bond JF, Garman RD, Griffith LJ, Kuo M, Pollock J;
 DR WPI: 94-035066/04.
 DE Antigen derived from Japanese cedar pollen allergen Cry j 1.
 PI Antigen derived from Japanese cedar pollen allergen Cry j 1.
 PI contain at least two T cell epitope(s), used to treat or diagnose
 PT allergy.
 PS Claim 76: Fig 18: 137pp: English.
 CC The sequence is that of an isolated peptide of the Japanese cedar
 CC pollen allergen Cry j 1. The peptide, CJI-43.7, can be used for
 CC the treatment and diagnosis of allergies associated with Japanese
 CC cedar pollen. It has enhanced therapeutic properties but reduced
 CC side effects compared to naturally occurring allergens.
 SQ Sequence 28 AA.

Query Match 100.0% Score 106; DB 1; Length 28;
 Best Local Similarity 100.0% Pred. No. 8.38e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DE 1 KSMKVTVAFNQFGPN 15
 QY 1 KSMKVTVAFNQFGPN 15

RESULT 13
 ID R45587 standard: Protein: 30 AA.

AC R45587:
 DT 13-JUL-1994 (first entry)
 DE Cry j 1 pollen allergen peptide CJI-43.1.
 KW Japanese cedar; detection: allergy; treatment: diagnosis;
 KW T cell epitope; sensitivity.
 OS Cryptomeria japonica.
 PN WO9401560-A.
 PD 20-JAN-1994.
 PF 15-JAN-1993; U00139.
 PR 10-JUL-1992: WO-005661.
 PA (IMMU-) IMMUNOLOGIC PHARM CORP.
 PI Bond JF, Garman RD, Griffith LJ, Kuo M, Pollock J;
 DR WPI: 94-035066/04.
 PI Antigen derived from Japanese cedar pollen allergen Cry j 1.
 PI contain at least two T cell epitope(s), used to treat or diagnose
 PT allergy.
 PS Claim 76: Fig 18: 137pp: English.
 CC The sequence is that of an isolated peptide of the Japanese cedar
 CC pollen allergen Cry j 1. The peptide, CJI-43.1, can be used for
 CC the treatment and diagnosis of allergies associated with Japanese
 CC cedar pollen. It has enhanced therapeutic properties but reduced
 CC side effects compared to naturally occurring allergens.
 SQ Sequence 30 AA.

Query Match 100.0% Score 106; DB 1; Length 30;
 Best Local Similarity 100.0% Pred. No. 8.38e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DE 1 KSMKVTVAFNQFGPN 15
 QY 1 KSMKVTVAFNQFGPN 15

RESULT 14

ID R45588 standard: Protein: 30 AA.

AC R45588:
 DT 13-JUL-1994 (first entry)
 DE Cry j 1 pollen allergen peptide CJI-43.6.
 KW Japanese cedar; detection: allergy; treatment: diagnosis;
 KW T cell epitope; sensitivity.
 OS Cryptomeria japonica.
 PN WO9401560-A.
 PD 20-JAN-1994.
 PF 15-JAN-1993; U00139.
 PR 10-JUL-1992: WO-005661.
 PR 01-SEP-1992: US-938990.

PA (IMMUNO) IMMUNOLOGIC PHARM CORP.
 PI Bond JP, Garzan RD, Griffith LJ, Kuo M, Collock J.
 PR WPI: 94-035065/04.
 PT Antigen derived from Japanese cedar pollen allergen Cry 1.
 PU Contains at least two T cell epitope(s), used to treat or diagnose
 PV allergy.
 PS Claim 76; Fig 18; 137pp; English.
 PP The sequence is that of an isolated peptide of the Japanese cedar
 PQ pollen allergen Cry 1. The peptide, CRI-43.6, can be used for
 PR the treatment and diagnosis of allergies associated with Japanese
 PS cedar pollen. It has enhanced therapeutic properties but reduced
 PT side effects compared to naturally occurring allergens.
 PU Sequence 30 AA:

Query Match 100.0% Score 100; 31; Length 30;
 Best Local Similarity 100.0%; Pred. No. 8, 18e-05;
 Matches 15; Conservative 0; Mismatches 0; Gaps 0;

LA 1 KSMKVIVAFNFGPN 15
 LV 1 KSMKV VAFNFGPN 15

RESULT 15
 ID W44587 standard: peptide: 30 AA.
 AC W44587.
 DT 01-MAY-1998 (first entry)
 DE T-cell epitope peptide #6 of sugi pollen antigen.
 KW T-cell epitope: sugi pollen antigen; sugi pollinosis.
 OS Synthetic.
 PS Crypomeria japonica
 PN J10007760-A.
 PD 13-JAN-1998.
 PE 24-JUN-1998; 163287.
 PP 24-JUN-1998; JP-163287.
 RA (DAIIC) DAIICHEM IND LTD.
 SA (MELB) MELB SEIKA KATSHA LTD.
 SW WPI: 98-035680/15.
 PT T cell epitope peptide of sugi pollen antigen - useful in the
 PU treatment of sugi pollinosis
 PS Claim 1; Page 4; 14pp; Japanese.
 PP T-cell epitope peptides W44582-88 and their derivatives react with
 PQ sugi pollinosis patient peripheral blood T lymphocytes. A composition
 PR prepared by combining at least 2 of the above peptides and/or their
 PS derivatives is used for the prevention and treatment of sugi pollinosis.
 PU Sequence 30 AA:

Query Match 100.0% Score 100; 31; Length 30;
 Best Local Similarity 100.0%; Pred. No. 8, 18e-05;
 Matches 15; Conservative 0; Mismatches 0; Gaps 0;

LA 1 KSMKVIVAFNFGPN 25
 LV 1 KSMKV VAFNFGPN 15

Search completed: Mon Jun 19 16:10:31 2000
 Job time: 11 secs

DB 256 KSMKVIVAFNQFGPN 15
LENGTH: 397 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE 397 AA: 42947 MW: 781425 CN:
Query Match 53.8% Score 57: DB 1: Length 397:
Best Local Similarity 60.0% Pred. No. 2.10e-01:
Matches 9: Conservative 1: Mismatches 5: Indels 0: Gaps 0:

DB 256 KSMKVIVAFNQFGPN 15
LENGTH: 397 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE 397 AA: 42947 MW: 781425 CN:
Query Match 53.8% Score 57: DB 1: Length 397:
Best Local Similarity 60.0% Pred. No. 2.10e-01:
Matches 9: Conservative 1: Mismatches 5: Indels 0: Gaps 0:

DB 256 KSMKVIVAFNQFGPN 15
LENGTH: 397 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE 397 AA: 42947 MW: 781425 CN:
Query Match 53.8% Score 57: DB 1: Length 397:
Best Local Similarity 60.0% Pred. No. 2.10e-01:
Matches 9: Conservative 1: Mismatches 5: Indels 0: Gaps 0:

DB 256 KSMKVIVAFNQFGPN 15
LENGTH: 397 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE 397 AA: 42947 MW: 781425 CN:
Query Match 53.8% Score 57: DB 1: Length 397:
Best Local Similarity 60.0% Pred. No. 2.10e-01:
Matches 9: Conservative 1: Mismatches 5: Indels 0: Gaps 0:

DB 256 KSMKVIVAFNQFGPN 15
LENGTH: 397 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE 397 AA: 42947 MW: 781425 CN:
Query Match 53.8% Score 57: DB 1: Length 397:
Best Local Similarity 60.0% Pred. No. 2.10e-01:
Matches 9: Conservative 1: Mismatches 5: Indels 0: Gaps 0:

DB 256 KSMKVIVAFNQFGPN 15
LENGTH: 397 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE 397 AA: 42947 MW: 781425 CN:
Query Match 53.8% Score 57: DB 1: Length 397:
Best Local Similarity 60.0% Pred. No. 2.10e-01:
Matches 9: Conservative 1: Mismatches 5: Indels 0: Gaps 0:

```
DE Sequence 59, Application US/08290448A
XX Sequence 59, Application US/08290448A
CC Patent No. 5676954
CC GENERAL INFORMATION:
CC APPLICANT: Rogers, Bruce
CC APPLICANT: Klapper, David G.
CC APPLICANT: Rafnar, Thorunn
CC APPLICANT: Kuo, Mei-chang
CC TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses
CC NUMBER OF SEQUENCES: 93
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: LAHIVE & COCKFIELD
CC STREET: 60 State Street, suite 510
CC CITY: Boston
CC STATE: Massachusetts
CC COUNTRY: USA
CC ZIP: 02109-1875
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/290.448A
CC FILING DATE: August 15, 1994
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/529,951
CC FILING DATE: May 29, 1990
CC APPLICATION NUMBER: US 07/325,365
CC FILING DATE: March 17, 1989
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Amy E. Mandragouras
CC REGISTRATION NUMBER: 36,207
CC REFERENCE/DOCKET NUMBER: IM-019CN
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617)227-7400
CC TELEFAX: (617)227-5941
CC INFORMATION FOR SEQ ID NO: 59:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 391 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 391 AA: 42854 MW 754045 CN:
CC
CC Query Match 51.9%, Score 55: DB 17, Length 391:
CC Best Local Similarity 66.7%, Pred. No. 3,436-01:
CC Matches 8: Conservative 1: Mismatches 3: Indels 0: Gaps 0:
CC
DB 250 KGMLATVAFNMF 261
QY 1 KSMKVTVAFNQF 12
CC
CC RESULT 5
CC US-08-290-448A-59 STANDARD: PRT: 391 AA.
CC
CC XXXXXX
CC
CC Sequence 59, Application US/08290448A
CC Sequence 59, Application US/08290448A
CC Patent No. 5698204
CC GENERAL INFORMATION:
CC APPLICANT: Rogers, Bruce
CC APPLICANT: Klapper, David G.
CC APPLICANT: Rafnar, Thorunn
CC APPLICANT: Kuo, Mei-chang
CC TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses
CC NUMBER OF SEQUENCES: 93
CC
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: LAHIVE & COCKFIELD, LLP
CC STREET: 60 State Street
CC CITY: Boston
CC STATE: Massachusetts
CC COUNTRY: USA
CC ZIP: 02109-1875
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: LAHIVE & COCKFIELD
CC STREET: 60 State Street
CC CITY: Boston
CC STATE: Massachusetts
CC COUNTRY: USA
CC ZIP: 02109-1875
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC
```

CC SOFTWARE: PatentIn Release #1 0. Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/175,069A
CC FILING DATE: December 29, 1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/529,951
CC FILING DATE: May 29, 1990
CC APPLICATION NUMBER: US 07/325,365
CC FILING DATE: March 17, 1989
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Amy E. Mandragouras
CC REGISTRATION NUMBER: 36,207
CC REFERENCE/DOCKET NUMBER: IM-018CN
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617)227-7400
CC TELEFAX: (617)227-5941
CC INFORMATION FOR SEQ ID NO: 59:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 391 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE: 391 AA: 42864 MW: 754045 CN:

Query Match 51.9% Score 55: DB 1: Length 391:
Best Local Similarity 66.7% Pred. NO. 3,438-01:
Matches 8: Conservative 1: Mismatches 3: Indels 0: Gaps 0:

DB 250 KGMALAIVAENMF 261
QY 1 KSMKVTVAENMF 12

RESULT 7
CC US-08-290-448A-74 STANDARD: PRI: 398 AA:
CC xxxxxx
CC
CC Sequence 74, Application US/08290448A
CC Patent No. 5676554
CC GENERAL INFORMATION:
CC APPLICANT: Rogers, Bruce
CC APPLICANT: Klapfer, David G.
CC APPLICANT: Rafnar, Thorunn
CC APPLICANT: Kuo, Mei-chang
CC TITLE OF INVENTION: Allergenic Proteins From Radweed and Uses
CC NUMBER OF SEQUENCES: 93
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: LAHIVE & COCKFIELD
CC STREET: 60 State Street, suite 510
CC CITY: Boston
CC STATE: Massachusetts
CC COUNTRY: USA
CC ZIP: 02109-1875
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/290,448A
CC FILING DATE: August 15, 1994
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/529,951
CC FILING DATE: May 29, 1990
CC APPLICATION NUMBER: US 07/325,365
CC FILING DATE: March 17, 1989
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Amy E. Mandragouras
CC REGISTRATION NUMBER: 36,207
CC REFERENCE/DOCKET NUMBER: IM-018CN
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617)227-7400
CC TELEFAX: (617)227-5941
CC INFORMATION FOR SEQ ID NO: 59:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 398 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein

CC REGISTRATION NUMBER: 36,207
CC REFERENCE/DOCKET NUMBER: IM-018CN
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617)227-7400
CC TELEFAX: (617)227-5941
CC INFORMATION FOR SEQ ID NO: 74:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 398 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE: 398 AA: 43564 MW: 739000 CN:

Query Match 51.9% Score 55: DB 1: Length 398:
Best Local Similarity 66.7% Pred. NO. 3,440-01:
Matches 8: Conservative 1: Mismatches 3: Indels 0: Gaps 0:

DB 257 KGMALAIVAENMF 258
QY 1 KSMKVTVAENMF 12

RESULT 8
CC US-08-290-448A-74 STANDARD: PRI: 498 AA:
CC xxxxxx
CC
CC Sequence 74, Application US/08290448A
CC Patent No. 5698204
CC GENERAL INFORMATION:
CC APPLICANT: Rogers, Bruce
CC APPLICANT: Klapfer, David G.
CC APPLICANT: Rafnar, Thorunn
CC APPLICANT: Kuo, Mei-chang
CC TITLE OF INVENTION: Allergenic Proteins From Radweed and Uses
CC NUMBER OF SEQUENCES: 93
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: LAHIVE & COCKFIELD
CC STREET: 60 State Street, suite 510
CC CITY: Boston
CC STATE: Massachusetts
CC COUNTRY: USA
CC ZIP: 02109-1875
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/290,448A
CC FILING DATE: August 15, 1994
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/529,951
CC FILING DATE: May 29, 1990
CC APPLICATION NUMBER: US 07/325,365
CC FILING DATE: March 17, 1989
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Amy E. Mandragouras
CC REGISTRATION NUMBER: 36,207
CC REFERENCE/DOCKET NUMBER: IM-018CN
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617)227-7400
CC TELEFAX: (617)227-5941
CC INFORMATION FOR SEQ ID NO: 74:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 398 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein

SEQUENCE 498 AA: 43664 MW: 779006 CN:

Query Match 51.9% Score 55: DB 1: Length: 399;
Best Local Similarity 66.7% Pred. No. 3.43e+01;
Matches 8: Conservative 1: Mismatches 3: Indels 0: Gaps 0:

DB 257 KGMLATVAFNMF 268
1:1 |||||
QY 1 KSMKVTVAFNQF 12

RESULT 9
ID US-08-175-069A-74 STANDARD: PR: 398 AA.
XX
AC xxxxxx
DE

Sequence 74. Application US/08175069A

Sequence 74. Application US/08175069A
Patent No. 5776761

GENERAL INFORMATION:

APPLICANT: Rogers, Bruce

APPLICANT: Klapper, David G.

APPLICANT: Rainar, Thorunn

APPLICANT: Kuo, Mei-chang

TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses

NUMBER OF SEQUENCES: 93

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD, LLP

STREET: 60 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109-1875

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version # 25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/175,069A

FILING DATE: December 29, 1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/529,951

FILING DATE: May 24, 1990

APPLICATION NUMBER: US/07/525,365

FILING DATE: March 17, 1989

ATTORNEY/AGENT INFORMATION:

NAME: Amy E. Mandragoras

REGISTRATION NUMBER: 36,207

REFERENCE/DOCKET NUMBER: IM-018DV

TELEPHONE: (617)227-7400

TELEFAX: (617)227-5941

INFORMATION FOR SEQ ID NO: 74:

SEQUENCE CHARACTERISTICS:

LENGTH: 398 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE 398 AA: 43664 MW: 779006 CN:

Query Match 51.9% Score 55: DB 1: Length: 398;
Best Local Similarity 66.7% Pred. No. 3.43e+01;
Matches 8: Conservative 1: Mismatches 3: Indels 0: Gaps 0:

DB 257 KGMLATVAFNMF 268
1:1 |||||
QY 1 KSMKVTVAFNQF 12

RESULT 10
ID US-08-290-448A-72 STANDARD: PR: 387 AA.
XX
AC xxxxxx
DE

Sequence 72. Application US/08290448A

Sequence 72. Application US/08290448A

Patent No. 5698204

GENERAL INFORMATION:

APPLICANT: Rogers, Bruce

APPLICANT: Klapper, David G.

APPLICANT: Rainar, Thorunn

APPLICANT: Kuo, Mei-chang

TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses

NUMBER OF SEQUENCES: 93

CORRESPONDENCE ADDRESS:

ADDRESSEE: LAHIVE & COCKFIELD

STREET: 60 State Street, suite 510

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109-1875

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/290,448A

FILING DATE: August 15, 1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/07/529,951

FILING DATE: May 29, 1990

APPLICATION NUMBER: US/07/325,365

FILING DATE: March 17, 1989

ATTORNEY/AGENT INFORMATION:

NAME: Amy E. Mandragoras

REGISTRATION NUMBER: 36,207

REFERENCE/DOCKET NUMBER: IM-018CN

TELEPHONE: (617)227-7400

TELEFAX: (617)227-5941

INFORMATION FOR SEQ ID NO: 72:

SEQUENCE CHARACTERISTICS:

LENGTH: 387 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE 387 AA: 41660 MW: 735967 CN:

Query Match 50.9% Score 54: DB 1: Length 387;

Best Local Similarity 53.3% Pred. No. 4.37e+01;

Matches 8: Conservative 2: Mismatches 5: Indels 0: Gaps 0:

DB 245 RGMLATVAFNMF 260

1:1 |||||

QY 1 KSMKVTVAFNMF 15

RESULT 11

ID US-08-290-448A-72 STANDARD: PR: 387 AA.

XX

AC xxxxxx

DE

XX

DT

XX

DE

XX

CC

CC

CC

CC


```

GENERAL INFORMATION:
APPLICANT: Rogers, Bruce
APPLICANT: Klapper, David G.
APPLICANT: Rafnar, Thorunn
APPLICANT: Kuo, Mei-chang
TITLE OF INVENTION: Allergenic Proteins From Raqweed and Uses
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/290.443A
FILING DATE: August 15, 1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/529,951
FILING DATE: May 29, 1990
APPLICATION NUMBER: US 07/325,365
FILING DATE: March 17, 1989
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragoras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMI-018CN
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE 387 AA: 41660 MW: 736967 CN:

Query Match 50.9% Score 54: DB 1 Length 187:
Best Local Similarity 53.3% Pred. No. 4,37e+01:
Matches 8: Conservative 2: Mismatches 5: Indels 0: Gaps 0:

E 246 RGMLAVAFNFIIDN 260
E 1 KSMKVIVAFNFIIDN 15
E 1 KSMKVIVAFNFIIDN 15

RESULT 12
US-08-175-069A-72 STANDARD: PRI: 387 AA.
xxxxxx
Sequence 72, Application US/08175069A
Patent No. 5775761
GENERAL INFORMATION:
APPLICANT: Rogers, Bruce
APPLICANT: Klapper, David G.
APPLICANT: Rafnar, Thorunn
APPLICANT: Kuo, Mei-chang
TITLE OF INVENTION: Allergenic Proteins From Raqweed and Uses
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 60 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/175,069A
FILING DATE: December 29, 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/529,951
FILING DATE: May 29, 1990
APPLICATION NUMBER: US 07/325,365
FILING DATE: March 17, 1989
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragoras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMI-018CN
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE 387 AA: 41660 MW: 736967 CN:

```

```

GENERAL INFORMATION:
APPLICANT: Rogers, Bruce
APPLICANT: Klapper, David G.
APPLICANT: Rafnar, Thorunn
APPLICANT: Kuo, Mei-chang
TITLE OF INVENTION: Allergenic Proteins From Raqweed and Uses
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 60 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/175,069A
FILING DATE: December 29, 1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/529,951
FILING DATE: May 29, 1990
APPLICATION NUMBER: US 07/325,365
FILING DATE: March 17, 1989
ATTORNEY/AGENT INFORMATION:
NAME: Amy E. Mandragoras
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: IMI-018DV
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE 387 AA: 41660 MW: 736967 CN:

Query Match 50.9% Score 54: DB 1 Length 187:
Best Local Similarity 53.3% Pred. No. 4,37e+01:
Matches 8: Conservative 2: Mismatches 5: Indels 0: Gaps 0:

DB 246 RGMLAVAFNFIIDN 260
CY 1 KSMKVIVAFNFIIDN 15

RESULT 13
US-08-175-069A-78 STANDARD: PRI: 383 AA.
xxxxxx
Sequence 78, Application US/08175069A
Patent No. 5775761
GENERAL INFORMATION:
APPLICANT: Rogers, Bruce
APPLICANT: Klapper, David G.
APPLICANT: Rafnar, Thorunn
APPLICANT: Kuo, Mei-chang
TITLE OF INVENTION: Allergenic Proteins From Raqweed and Uses
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD, LLP
STREET: 60 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/175,069A
FILING DATE: December 29, 1993

```

CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: US 07/529,951
 CC FILING DATE: May 29, 1990
 CC APPLICATION NUMBER: US 07/325,365
 CC FILING DATE: March 17, 1989
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Amy E. Mandragoras
 CC REGISTRATION NUMBER: 36,207
 CC REFERENCE/DOCKET NUMBER: IMI-0182V
 CC TELEPHONE: (617)227-7400
 CC TELEFAX: (617)227-5441
 CC INFORMATION FOR SEQ ID NO: 78:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 383 amino acids
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC SEQUENCE 383 AA: 41852 MW: 738971 CN:

Query Match 50.0% Score 53: DB 1: Length 383:
 Best Local Similarity 53.3% Pred. No. 5.57e+01:
 Matches 8: Conservative 2: Mismatches 5: Indels 0: Gaps 0:

DE 243 RGMICVAFNKFIDN 257
 QY 1 KSMKVIVAFNQFSPN 15

RESULT 14
 ID US-08-290-448A-78 STANDARD: PRT: 383 AA.

AC xxxxxx

Sequence 78, Application US/08290448A

Sequence 78, Application US/08290448A
 Patent No. 5676954

CC GENERAL INFORMATION:
 CC APPLICANT: Rogers, Bruce
 CC APPLICANT: Klapper, David G.
 CC APPLICANT: Rainar, Thorunn
 CC APPLICANT: Kuo, Mei-Chang
 CC TITLE OF INVENTION: Antigenic Proteins From Kaposar and Uses
 CC NUMBER OF SEQUENCES: 93
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: LAHIVE & COCKFIELD
 CC STREET: 60 State Street, Suite 510
 CC CITY: Boston
 CC STATE: Massachusetts
 CC COUNTRY: USA
 CC ZIP: 02109-1875

CC COMPUTER READABLE FORM:
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patent Release #1.0, Version #1.25

CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/290.448A
 CC FILING DATE: August 15, 1994

CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: US 07/529,951
 CC FILING DATE: May 29, 1990

CC APPLICATION NUMBER: US 07/325,365
 CC FILING DATE: March 17, 1989

CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Amy E. Mandragoras
 CC REGISTRATION NUMBER: 36,207
 CC REFERENCE/DOCKET NUMBER: IMI-018CN
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (617)227-7400

CC TELEFAX: (617)227-5941
 CC INFORMATION FOR SEQ ID NO: 78:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 383 amino acids
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC SEQUENCE 383 AA: 41852 MW: 738971 CN:

Query Match 50.0% Score 53: DB 1: Length 383:
 Best Local Similarity 53.3% Pred. No. 5.57e+01:
 Matches 8: Conservative 2: Mismatches 5: Indels 0: Gaps 0:

DE 243 RGMICVAFNKFIDN 257
 QY 1 KSMKVIVAFNQFSPN 15

RESULT 15
 ID US-08-290-448A-78 STANDARD: PRT: 383 AA

AC xxxxxx

Sequence 78, Application US/08290448A

Sequence 78, Application US/08290448A
 Patent No. 5698204

CC GENERAL INFORMATION:
 CC APPLICANT: Rogers, Bruce
 CC APPLICANT: Klapper, David G.
 CC APPLICANT: Rainar, Thorunn
 CC APPLICANT: Kuo, Mei-Chang
 CC TITLE OF INVENTION: Antigenic Proteins From Kaposar and Uses
 CC NUMBER OF SEQUENCES: 93
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: LAHIVE & COCKFIELD
 CC STREET: 60 State Street, Suite 510
 CC CITY: Boston
 CC STATE: Massachusetts
 CC COUNTRY: USA
 CC ZIP: 02109-1875

CC COMPUTER READABLE FORM:
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patent Release #1.0, Version #1.25

CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/290.448A
 CC FILING DATE: August 15, 1994

CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: US 07/529,951
 CC FILING DATE: May 29, 1990

CC APPLICATION NUMBER: US 07/325,365
 CC FILING DATE: March 17, 1989

CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Amy E. Mandragoras
 CC REGISTRATION NUMBER: 36,207
 CC REFERENCE/DOCKET NUMBER: IMI-018CN
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (617)227-7400

CC TELEFAX: (617)227-5441
 CC INFORMATION FOR SEQ ID NO: 78:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 383 amino acids
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC SEQUENCE 383 AA: 41852 MW: 738971 CN:

Query Match 50.0% Score 53: DB 1: Length 383:
 Best Local Similarity 53.3% Pred. No. 5.57e+01:

Matches 8: Conservative 2: Mismatches 5: Indels 0: Gaps 0:

PI 243 RSMICTVAFNKKFTDN 257
III III I I
Q3 1 KSMKVTVAFNLCGN 15

Search completed: Tue Jun 20 13:46:01 2000
Job time : 5 secs.

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPSrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Mon Jun 19 16:26:01 2000; MasPar time 14.94 Seconds
Tabular output not generated.
Title: >US-09-142-524A-9
Description: (1-15) from US09142524A.pep
Sequence: 1 KSMKVTVAFNQFGPN 15
Scoring table: PAM 150
Gap 15
Searched: 721208 seqs. 100765575 residues
Post-processing: Minimum Match 0%
Listing first 45 summaries
Database: a-pending
1: PCT 2:06 3:060 4:07 5:06 5:081 7:082 8:083 9:084
10:084B 11:085 12:086 13:07 14:089 15:089 16:090 17:091
18:092 19:093 20:094 21:05 22:NEWB 23:NEWU60 24:NEWU8
25:NEWU9
Statistics: Mean 20.547 Variance 53.151; scale 0.366
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.
SUMMARIES
Result No. Query Match Length DB ID Description Pred. No.
1 106 100.0 19 10 US-08-467- Sequence 227, Applicat 4.81e-05
2 106 100.0 19 8 US-08-350- Sequence 227, Applicat 4.81e-05
3 106 100.0 20 10 US-08-467- Sequence 230, Applicat 4.81e-05
4 106 100.0 20 4 US-07-938- Sequence 47, Applicati 4.81e-05
5 106 100.0 20 8 US-08-350- Sequence 230, Applicat 4.81e-05
6 106 100.0 20 10 US-08-467- Sequence 47, Applicati 4.81e-05
7 106 100.0 21 10 US-08-467- Sequence 231, Applicat 4.81e-05
8 106 100.0 21 10 US-08-467- Sequence 231, Applicat 4.81e-05
9 106 100.0 23 10 US-08-468- Sequence 224, Applicat 4.81e-05
10 106 100.0 23 10 US-08-468- Sequence 224, Applicat 4.81e-05
11 106 100.0 26 10 US-08-468- Sequence 86, Applicati 4.81e-05
12 106 100.0 26 10 US-08-468- Sequence 83, Applicati 4.81e-05
13 106 100.0 26 10 US-08-468- Sequence 83, Applicati 4.81e-05
14 106 100.0 26 10 US-08-467- Sequence 86, Applicati 4.81e-05
15 106 100.0 26 10 US-08-467- Sequence 83, Applicati 4.81e-05
16 106 100.0 26 10 US-08-467- Sequence 84, Applicati 4.81e-05
17 106 100.0 26 10 US-08-467- Sequence 83, Applicati 4.81e-05
18 106 100.0 28 10 US-08-468- Sequence 81, Applicati 4.81e-05
19 106 100.0 28 10 US-08-467- Sequence 82, Applicati 4.81e-05
20 106 100.0 28 8 US-08-350- Sequence 81, Applicati 4.81e-05

21 106 100.0 28 10 US-08-467- Sequence 81, Applicat 4.81e-05
22 106 100.0 28 8 US-08-150- Sequence 219, Applicat 4.81e-05
23 106 100.0 28 10 US-08-468- Sequence 85, Applicat 4.81e-05
24 106 100.0 28 10 US-08-467- Sequence 82, Applicat 4.81e-05
25 106 100.0 28 10 US-08-467- Sequence 85, Applicat 4.81e-05
26 106 100.0 28 7 US-08-226- Sequence 82, Applicat 4.81e-05
27 106 100.0 28 10 US-08-468- Sequence 82, Applicat 4.81e-05
28 106 100.0 28 10 US-08-467- Sequence 81, Applicat 4.81e-05
29 106 100.0 28 10 US-08-467- Sequence 219, Applicat 4.81e-05
30 106 100.0 28 10 US-08-467- Sequence 82, Applicat 4.81e-05
31 106 100.0 29 10 US-08-467- Sequence 222, Applicat 4.81e-05
32 106 100.0 30 10 US-08-467- Sequence 79, Applicat 4.81e-05
33 106 100.0 30 7 US-08-226- Sequence 79, Applicat 4.81e-05
34 106 100.0 30 10 US-08-467- Sequence 80, Applicat 4.81e-05
35 106 100.0 30 10 US-08-468- Sequence 80, Applicat 4.81e-05
36 106 100.0 30 10 US-08-467- Sequence 83, Applicat 4.81e-05
37 106 100.0 36 10 US-08-467- Sequence 78, Applicat 4.81e-05
38 106 100.0 36 10 US-08-467- Sequence 78, Applicat 4.81e-05
39 106 100.0 90 10 US-08-468- Sequence 54, Applicat 4.81e-05
40 106 100.0 90 4 US-07-938- Sequence 54, Applicat 4.81e-05
41 106 100.0 90 10 US-08-467- Sequence 54, Applicat 4.81e-05
42 106 100.0 367 10 US-08-467- Sequence 95, Applicat 4.81e-05
43 106 100.0 367 10 US-08-467- Sequence 95, Applicat 4.81e-05
44 106 100.0 374 8 US-08-350- Sequence 2, Applicatio 4.81e-05
45 106 100.0 374 4 US-07-938- Sequence 2, Applicatio 4.81e-05

ALIGNMENTS

RESULT 1
ID US-08-467-023-227 STANDARD: PRT: 19 AA
XX
AC xxxxxx
XX
DT
XX
DE Sequence 227, Application US/08467023
XX
CC Sequence 227, Application US/08467023
CC GENERAL INFORMATION:
CC APPLICANT: Griffith, Irwin J.
CC APPLICANT: Pollock, Joanne.
CC APPLICANT: Reed, Julian P.
CC APPLICANT: Gaiman, Richard D.
CC APPLICANT: Koo, Melvina.
CC APPLICANT: Yeh, Sigmei H.
CC APPLICANT: Brauer, Andrew.
CC APPLICANT: Exley, Mark A.
CC APPLICANT: Powels, Steven F.
CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
CC TITLE OF INVENTION: Japanese Cedar Pollen
CC NUMBER OF SEQUENCES: 261
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
CC STREET: 610 Lincoln St
CC CITY: Waltham
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02154
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION NUMBER: US/08/467.023
CC FILING DATE: June 6, 1995
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/350,225
CC FILING DATE: December 6, 1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Jane E. Remilla-3

CC REGISTRATION NUMBER: 38,872
 CC REFERENCE/DOCKET NUMBER: 025.6 USD2 (IM1-028CPD2)
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (617) 466-6000
 CC TELEFAX: (617) 466-6040
 CC INFORMATION FOR SEQ ID NO: 227:
 CC TELEFAX: (617) 227-7400
 CC INFORMATION FOR SEQ ID NO: 227:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 19 amino acids
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: peptide
 CC FRAGMENT TYPE: internal
 CC M-SEQUENCE TYPE: internal
 CC SEQUENCE 19 AA: 2156 MW: 1801 CN:
 CC
 CC Query Match 100.0% Score 106.0 Length 19:
 CC Best Local Similarity 100.0% Pred. No. 4.8e-05:
 CC Matches 15: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 CC
 CC Db 3 KSMKVIVAFNFGPN 17
 CC 1 1111111111
 CC QY 1 KSMKVIVAFNFGPN 15
 CC
 CC RESULT 2
 CC ID US-09-350-225-227 STANDARD: PRI: 19 AA:
 CC XX xxxxxx
 CC
 CC Sequence 227, Application US/08350225
 CC Sequence 227, Application US/08350225
 CC GENERAL INFORMATION:
 CC APPLICANT: Griffith, Irwin J.
 CC APPLICANT: Pollock, Joanne
 CC APPLICANT: Bond, Julian F.
 CC APPLICANT: Gattac, Richard D.
 CC APPLICANT: Kuo, Mei-Chang
 CC APPLICANT: Yeung, Siu-mei H.
 CC APPLICANT: Brauer, Andrew
 CC APPLICANT: Exley, Mark A.
 CC APPLICANT: Powers, Steven P.
 CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
 CC TITLE OF INVENTION: Japanese Cedar Pollen
 CC NUMBER OF SEQUENCES: 2
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Immunogen Pharmaceuticals Corporation, Inc.
 CC STREET: 610 Lincoln St.
 CC CITY: Waltham
 CC STATE: MA
 CC COUNTRY: USA
 CC ZIP: 02154
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: PatentIn Release #1.0, Version #1.05
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/350,225
 CC FILING DATE: December 6, 1994
 CC CLASSIFICATION: 424
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/226,248
 CC FILING DATE: April 8, 1994
 CC APPLICATION NUMBER: 07/948,990
 CC FILING DATE: September 1, 1992
 CC APPLICATION NUMBER: PCT/US93/00139
 CC FILING DATE: January 15, 1993
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Carlene A. Varsone
 CC REGISTRATION NUMBER: 35,729
 CC REFERENCE/DOCKET NUMBER: 025.6 US (IM1-028CP2)

CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (617) 466-6000
 CC TELEFAX: (617) 466-6040
 CC INFORMATION FOR SEQ ID NO: 227:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 19 amino acids
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: peptide
 CC FRAGMENT TYPE: internal
 CC SEQUENCE 19 AA: 2156 MW: 1801 CN:
 CC
 CC Query Match 100.0% Score 106.0 Length 19:
 CC Best Local Similarity 100.0% Pred. No. 4.8e-05:
 CC Matches 15: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 CC
 CC Db 3 KSMKVIVAFNFGPN 17
 CC 1 1111111111
 CC QY 1 KSMKVIVAFNFGPN 15
 CC
 CC RESULT 3
 CC ID US-09-467-023-230 STANDARD: PRI: 20 AA:
 CC XX xxxxxx
 CC
 CC Sequence 210, Application US/08467225
 CC Sequence 210, Application US/08467225
 CC GENERAL INFORMATION:
 CC APPLICANT: Griffith, Irwin J.
 CC APPLICANT: Pollock, Joanne
 CC APPLICANT: Bond, Julian F.
 CC APPLICANT: Gattac, Richard D.
 CC APPLICANT: Kuo, Mei-Chang
 CC APPLICANT: Yeung, Siu-mei H.
 CC APPLICANT: Brauer, Andrew
 CC APPLICANT: Exley, Mark A.
 CC APPLICANT: Powers, Steven P.
 CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
 CC TITLE OF INVENTION: Japanese Cedar Pollen
 CC NUMBER OF SEQUENCES: 241
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Immunogen Pharmaceuticals Corporation, Inc.
 CC STREET: 610 Lincoln St.
 CC CITY: Waltham
 CC STATE: MA
 CC COUNTRY: USA
 CC ZIP: 02154
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: PatentIn Release #1.0, Version #1.05
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/467,023
 CC FILING DATE: June 6, 1995
 CC CLASSIFICATION: 424
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/350,225
 CC FILING DATE: December 6, 1994
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Jane E. Remillard
 CC REGISTRATION NUMBER: 38,872
 CC REFERENCE/DOCKET NUMBER: 025.6 USD2 (IM1-028CP2)
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (617) 227-7400
 CC TELEFAX: (617) 227-5941
 CC INFORMATION FOR SEQ ID NO: 230:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 20 amino acids

CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC FRAGMENT TYPE: internal
SQ SEQUENCE 20 AA: 2285 MW: 2007 CN:

Query Match 100.0% Score 106; DB 10; Length 20;
Best Local Similarity 100.0% Pred. No. 4.81e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 4 KSMKVTVAFNQFGPN 18
QY 1 KSMKVTVAFNQFGPN 15

RESULT 4
ID US-07-938-953A-47 STANDARD: PRT: 20 AA.

XX XXXXXX

Sequence 47, Application US/07938990A

Sequence 47, Application US/07938990A

GENERAL INFORMATION:

APPLICANT: Griffith, Irwin J.

APPLICANT: Pollock, Joanne

APPLICANT: Bond, Julian

TITLE OF INVENTION: Allergenic Proteins And Peptides From

TITLE OF INVENTION: Japanese Cedar Pollen

NUMBER OF SEQUENCES: 70

CORRESPONDENCE ADDRESS:

ADDRESSES: Lahive & Cockfield

STREET: Sixty State Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/948,990A

FILING DATE: 1992-9-01

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/730,452

FILING DATE: July 15, 1991

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/739,134

FILING DATE: July 10, 1991

ATTORNEY/AGENT INFORMATION:

NAME: Amy E. Mandragoras

REGISTRATION NUMBER: 36,207

REFERENCE/DOCKET NUMBER: IPC-C25CC (IMI-028)

TELEPHONE: (617) 227-7400

TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 47:

SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: peptide

FRAGMENT TYPE: internal

SEQUENCE 20 AA: 2244 MW: 1852 CN:

Query Match 100.0% Score 106; DB 4; Length 20;
Best Local Similarity 100.0% Pred. No. 4.81e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 KSMKVTVAFNQFGPN 15
QY 1 KSMKVTVAFNQFGPN 15

RESULT 5
ID US-08-350-225-230 STANDARD: PRT: 20 AA.

XX XXXXXX

Sequence 230, Application US/08350225

Sequence 230, Application US/08350225

GENERAL INFORMATION:

APPLICANT: Griffith, Irwin J.

APPLICANT: Pollock, Joanne

APPLICANT: Bond, Julian

APPLICANT: Garman, Richard D.

APPLICANT: Kuo, Mei-Chang

APPLICANT: Yeung, Siu-mei H.

APPLICANT: Brauer, Andrew

APPLICANT: Exley, Mark A.

APPLICANT: Powers, Steven P.

TITLE OF INVENTION: Allergenic Proteins And Peptides From

TITLE OF INVENTION: Japanese Cedar Pollen

NUMBER OF SEQUENCES: 261

CORRESPONDENCE ADDRESS:

ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.

STREET: 510 Lincoln St

CITY: Waltham

STATE: MA

COUNTRY: USA

ZIP: 02154

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/350,225

FILING DATE: December 1, 1994

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/221,248

FILING DATE: April 8, 1994

APPLICATION NUMBER: 07/945,990

FILING DATE: September 1, 1992

APPLICATION NUMBER: PCI/US93/00139

FILING DATE: January 15, 1993

ATTORNEY/AGENT INFORMATION:

NAME: Darlene A. Vanstone

REGISTRATION NUMBER: 35,729

REFERENCE/DOCKET NUMBER: 025.6 US (LM-028CP2)

TELEPHONE: (617) 466-6700

TELEFAX: (617) 466-604

INFORMATION FOR SEQ ID NO: 230:

SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FRAGMENT TYPE: internal

SEQUENCE 20 AA: 2285 MW: 2007 CN:

Query Match 100.0% Score 106; DB 8; Length 20;
Best Local Similarity 100.0% Pred. No. 4.81e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 4 KSMKVTVAFNQFGPN 18

1 KSMKVTVAEFGPN 15

US-08-467-697-47 STANDARD: PRI: 20 AA.

xxxxxx

Sequence 47, Application US/08457697

Sequence 47, Application US/08457697

GENERAL INFORMATION:
 APPLICANT: Griffith, Irwin J.
 APPLICANT: Pollock, Joanne
 APPLICANT: Bond, Julia F.
 APPLICANT: Garman, Richard D.
 APPLICANT: Kuo, Mei-Chang
 APPLICANT: Yeung, Siu-mei H.
 APPLICANT: Brauer, Andrew
 APPLICANT: Exley, Mark A.
 APPLICANT: Powers, Steven P.
 TITLE OF INVENTION: Allergenic Proteins And Peptides From
 TITLE OF INVENTION: Japanese Cedar Pollen
 NUMBER OF SEQUENCES: 261
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
 STREET: 610 Lincoln St
 CITY: Waltham
 STATE: MA
 COUNTRY: USA
 ZIP: 02154

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION NUMBER: US/08/467,697
 FILING DATE: June 6, 1995
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/350,225
 FILING DATE: December 6, 1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Jane E. Remillard
 REGISTRATION NUMBER: 98,972
 REFERENCE/DOCKET NUMBER: 025.6 US4 (INT-0282704)
 TELEPHONE: (617) 227-7400
 TELEFAX: (617) 227-5941
 INFORMATION FOR SEQ ID NO: 47:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 20 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 FRAGMENT TYPE: internal

SEQUENCE 20 AA: 2244 MW: 1852 CN:

Query Match 100.0% Score 106; DB 10; Length 20;

Best Local Similarity 100.0% Pred. No. 4,81e-05;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 KSMKVTVAEFGPN 15

1 KSMKVTVAEFGPN 15

US-08-467-697-47

STANDARD: PRI: 21 AA.

xxxxxx

Sequence 47, Application US/08457697

Sequence 47, Application US/08457697

GENERAL INFORMATION:
 APPLICANT: Griffith, Irwin J.
 APPLICANT: Pollock, Joanne
 APPLICANT: Bond, Julia F.
 APPLICANT: Garman, Richard D.
 APPLICANT: Kuo, Mei-Chang
 APPLICANT: Yeung, Siu-mei H.
 APPLICANT: Brauer, Andrew
 APPLICANT: Exley, Mark A.
 APPLICANT: Powers, Steven P.
 TITLE OF INVENTION: Allergenic Proteins And Peptides From
 TITLE OF INVENTION: Japanese Cedar Pollen
 NUMBER OF SEQUENCES: 261
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
 STREET: 610 Lincoln St
 CITY: Waltham
 STATE: MA
 COUNTRY: USA
 ZIP: 02154

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION NUMBER: US/08/467,697
 FILING DATE: June 6, 1995
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/350,225
 FILING DATE: December 6, 1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Jane E. Remillard
 REGISTRATION NUMBER: 98,972
 REFERENCE/DOCKET NUMBER: 025.6 US4 (INT-0282704)
 TELEPHONE: (617) 227-7400
 TELEFAX: (617) 227-5941
 INFORMATION FOR SEQ ID NO: 47:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 20 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 FRAGMENT TYPE: internal

xxxxxx

Sequence 231, Application US/08467023

Sequence 231, Application US/08467023

GENERAL INFORMATION:
 APPLICANT: Griffith, Irwin J.
 APPLICANT: Pollock, Joanne
 APPLICANT: Bond, Julia F.
 APPLICANT: Garman, Richard D.
 APPLICANT: Kuo, Mei-Chang
 APPLICANT: Yeung, Siu-mei H.
 APPLICANT: Brauer, Andrew
 APPLICANT: Exley, Mark A.
 APPLICANT: Powers, Steven P.
 TITLE OF INVENTION: Allergenic Proteins And Peptides From
 TITLE OF INVENTION: Japanese Cedar Pollen
 NUMBER OF SEQUENCES: 261
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
 STREET: 610 Lincoln St
 CITY: Waltham
 STATE: MA
 COUNTRY: USA
 ZIP: 02154

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION NUMBER: US/08/467,023
 FILING DATE: June 6, 1995
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/350,225
 FILING DATE: December 6, 1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Jane E. Remillard
 REGISTRATION NUMBER: 98,972
 REFERENCE/DOCKET NUMBER: 025.6 US4 (INT-0282704)
 TELEPHONE: (617) 227-7400
 TELEFAX: (617) 227-5941
 INFORMATION FOR SEQ ID NO: 231:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 21 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 FRAGMENT TYPE: internal

SEQUENCE 21 AA: 2415 MW: 2154 CN:

Query Match 100.0% Score 106; DB 10; Length 21;

Best Local Similarity 100.0% Pred. No. 4,81e-05;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

4 KSMKVTVAEFGPN 18

1 KSMKVTVAEFGPN 15

US-08-467-006-231

STANDARD: PRI: 21 AA.

xxxxxx

Sequence 231, Application US/08467006

XX Sequence 231, Application US/08457006
 CC GENERAL INFORMATION:
 CC APPLICANT: Griffith, Irwin J.
 CC APPLICANT: Pollock, Joanne
 CC APPLICANT: Bond, Julian F.
 CC APPLICANT: Garman, Richard D.
 CC APPLICANT: Kuo, Mei-Cha
 CC APPLICANT: Yeung, Siu-mei H.
 CC APPLICANT: Brauer, Andrew
 CC APPLICANT: Exley, Mark A.
 CC APPLICANT: Powers, Steven P.
 CC APPLICANT: Griffith, Irwin J.
 CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
 CC TITLE OF INVENTION: Japanese Cedar Pollen
 CC NUMBER OF SEQUENCES: 26
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
 CC STREET: 610 Lincoln St
 CC CITY: Waltham
 CC STATE: MA
 CC COUNTRY: USA
 CC ZIP: 02154
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patent Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08457.006
 CC FILING DATE: June 6, 1995
 CC CLASSIFICATION: 424
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/350,225
 CC FILING DATE: December 6, 1994
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Zace E. Remillard
 CC REGISTRATION NUMBER: 38,872
 CC REFERENCE/DOCKET NUMBER: 025.6 US 3 (IM:028CPD5)
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (617) 227-7400
 CC TELEFAX: (617) 227-5941
 CC INFORMATION FOR SEQ ID NO: 231:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 21 amino acids
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: peptide
 CC FRAGMENT TYPE: internal
 CC SEQUENCE 21 AA: 2415 KW: 2154 CN:

Query Match 100.0% Score 106: DB 10: Length 21:
 Best Local Similarity 100.0% Pred. No. 4.81e-05:
 Matches 15: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

DI 4 KSMKVTVAFNQFGPN 18
 QY 1 KSMKVTVAFNQFGPN 15

RESULT 9
 ID US-08-468-940-233 STANDARD: PRI: 23 AA.
 XX
 AC xxxxxx
 XX
 DT

XX Sequence 233, Application US/08468940
 CC Sequence 233, Application US/08468940
 CC GENERAL INFORMATION:
 CC APPLICANT: Griffith, Irwin J.
 CC APPLICANT: Pollock, Joanne
 CC APPLICANT: Bond, Julian F.

CC APPLICANT: Garman, Richard
 CC APPLICANT: Kuo, Mei-Cha
 CC APPLICANT: Yeung, Siu-mei H.
 CC APPLICANT: Brauer, Andrew
 CC APPLICANT: Exley, Mark A.
 CC APPLICANT: Powers, Steven P.
 CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
 CC TITLE OF INVENTION: Japanese Cedar Pollen
 CC NUMBER OF SEQUENCES: 26
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
 CC STREET: 610 Lincoln St
 CC CITY: Waltham
 CC STATE: MA
 CC COUNTRY: USA
 CC ZIP: 02154
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patent Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08458.940
 CC FILING DATE:
 CC CLASSIFICATION: 424
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/350,225
 CC FILING DATE: December 6, 1994
 CC APPLICATION NUMBER: 08/226,248
 CC FILING DATE: April 8, 1994
 CC APPLICATION NUMBER: 07/936,990
 CC FILING DATE: September 1, 1992
 CC APPLICATION NUMBER: PCT/US93/00134
 CC FILING DATE: January 15, 1993
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Darlene A. Vanstone
 CC REGISTRATION NUMBER: 35,729
 CC REFERENCE/DOCKET NUMBER: 025.6 US (IM:028CP2)
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (617) 466-6040
 CC TELEFAX: (617) 466-6040
 CC INFORMATION FOR SEQ ID NO: 233:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 23 amino acids
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: peptide
 CC FRAGMENT TYPE: internal
 CC SEQUENCE 23 AA: 2558 MW: 2164 IN:

Query Match 100.0% Score 106: DB 10: Length 23:
 Best Local Similarity 100.0% Pred. No. 4.81e-05:
 Matches 15: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Db 4 KSMKVTVAFNQFGPN 18
 QY 1 KSMKVTVAFNQFGPN 15

RESULT 10
 ID US-08-467-023-224 STANDARD: PRI: 24 AA.
 XX
 AC xxxxxx
 XX
 DT

XX Sequence 224, Application US/08467023
 CC Sequence 224, Application US/08467023
 CC GENERAL INFORMATION:
 CC APPLICANT: Griffith, Irwin J.
 CC APPLICANT: Pollock, Joanne
 CC APPLICANT: Bond, Julian F.

CC APPLICANT: Garman, Richard D.
 CC APPLICANT: Kuo, Mei-Chang.
 CC APPLICANT: Yeung, Siu-mei H.
 CC APPLICANT: Brauer, Andrew.
 CC APPLICANT: Exley, Mark A.
 CC APPLICANT: Powers, Steven P.
 CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
 CC TITLE OF INVENTION: Japanese Cedar Pollen
 CC NUMBER OF SEQUENCES: 26.
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
 CC STREET: 610 Lincoln St
 CC CITY: Waltham
 CC STATE: MA
 CC COUNTRY: USA
 CC ZIP: 02154
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: PatentIn Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/467,023
 CC FILING DATE: June 6, 1995
 CC CLASSIFICATION: 424
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/350,225
 CC FILING DATE: December 5, 1994
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Jace E. Renillard
 CC REGISTRATION NUMBER: 38,872
 CC REFERENCE/DOCKET NUMBER: 025.6 US22 (IMI-028CPD2)
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (617) 227-7400
 CC TELEFAX: (617) 227-5941
 CC INFORMATION FOR SEQ ID NO: 224:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 24 amino acids
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: peptide
 CC FRAGMENT TYPE: internal
 CC SEQUENCE 24 AA: 2716 MW: 2744 CN:

Query Match 100.0% Score 105. DB 10 Length 24
 Best Local Similarity 100.0% Pred. No. 4.81e-05
 Matches 15: Conservative 83 Mismatches 00 Gaps 00

CL 3 KSKKVTVAFFNGPGN 17
 QY 1 KSKKVTVAFFNGPGN 15

RESULT 11
 ID US-08-468-940-86 STANDARD: PRI: 26 AA.
 XX AC xxxxxx
 XX DT
 XX DE

Sequence 86, Application US/08458940
 GENERAL INFORMATION:
 CC APPLICANT: Griffith, Irwin J.
 CC APPLICANT: Pollock, Joanne
 CC APPLICANT: Bond, Julian F.
 CC APPLICANT: Garman, Richard D.
 CC APPLICANT: Kuo, Mei-Chang
 CC APPLICANT: Yeung, Siu-mei H.
 CC APPLICANT: Brauer, Andrew
 CC APPLICANT: Exley, Mark A.
 CC APPLICANT: Powers, Steven P.

CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
 CC TITLE OF INVENTION: Japanese Cedar Pollen
 CC NUMBER OF SEQUENCES: 261
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
 CC STREET: 610 Lincoln St
 CC CITY: Waltham
 CC STATE: MA
 CC COUNTRY: USA
 CC ZIP: 02154
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: PatentIn Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/458,940
 CC FILING DATE:
 CC CLASSIFICATION: 424
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/350,225
 CC FILING DATE: December 5, 1994
 CC APPLICATION NUMBER: 08/226,248
 CC FILING DATE: April 8, 1994
 CC APPLICATION NUMBER: 07/48,990
 CC FILING DATE: September 1, 1992
 CC APPLICATION NUMBER: 01 US93/00139
 CC FILING DATE: January 15, 1993
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Darlene A. Vanstone
 CC REGISTRATION NUMBER: 35,729
 CC REFERENCE/DOCKET NUMBER: 025.6 US (IMI-028CP2)
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (617) 466-6400
 CC TELEFAX: (617) 466-8040
 CC INFORMATION FOR SEQ ID NO: 86:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 26 amino acids
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: peptide
 CC FRAGMENT TYPE: internal
 CC SEQUENCE 26 AA: 2944 MW: 2941 CN:

Query Match 100.0% Score 105. DB 10 Length 26
 Best Local Similarity 100.0% Pred. No. 1.4e-06
 Matches 15: Conservative 83 Mismatches 00 Gaps 00

CL 1 KSKKVTVAFFNGPGN 15
 QY 1 KSKKVTVAFFNGPGN 15

RESULT 12
 ID US-08-468-940-83 STANDARD: PRI: 26 AA.
 XX AC xxxxxx
 XX DT
 XX DE

Sequence 83, Application US/08458940
 GENERAL INFORMATION:
 CC APPLICANT: Griffith, Irwin J.
 CC APPLICANT: Pollock, Joanne
 CC APPLICANT: Bond, Julian F.
 CC APPLICANT: Garman, Richard D.
 CC APPLICANT: Kuo, Mei-Chang
 CC APPLICANT: Yeung, Siu-mei H.
 CC APPLICANT: Brauer, Andrew
 CC APPLICANT: Exley, Mark A.
 CC APPLICANT: Powers, Steven P.

CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
CC TITLE OF INVENTION: Japanese Cedar Pollen
CC NUMBER OF SEQUENCES: 251
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
CC STREET: 610 Lincoln St
CC CITY: Waltham
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02154
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/468,940
CC FILING DATE:
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/350,225
CC FILING DATE: December 6, 1994
CC APPLICATION NUMBER: 08/226,248
CC FILING DATE: April 8, 1994
CC APPLICATION NUMBER: 07/338,990
CC FILING DATE: September 11, 1992
CC APPLICATION NUMBER: PC7/US93/00139
CC FILING DATE: January 15, 1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Darlene A. Vanstone
CC REGISTRATION NUMBER: 35,729
CC REFERENCE/DOCKET NUMBER: 025.6 US (IMI-028C22)
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 466-6000
CC TELEFAX: (617) 466-6040
CC INFORMATION FOR SEQ ID NO: 83:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 26 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC FRAGMENT TYPE: internal
CC SEQUENCE 26 AA: 2944 MW: 2965 CN
SQ
Query Match 100.0% Score 104.00 Length 26
Best Local Similarity 100.0% Pred. No. 4,81e-05
Matches 157 Conservative 0 Mismatches 0 Indels 0 Gaps 0
DB 1 KSMKVTVAENQFNP 15
Q7 1 KSMKVTVAENQFNP 15
RESULT 13
ID US-08-226-248A-83 STANDARD: PRI: 25 AA:
XX xxxxxx
AC
XX
DT
DE
Sequence 83, Application US/08226248A
Sequence 83, Application US/08226248A
GENERAL INFORMATION:
CC APPLICANT: Griffith, Irwin J.:
CC APPLICANT: Pollock, Joanne:
CC APPLICANT: Bond, Julian F.:
CC APPLICANT: Garman, Richard D.:
CC APPLICANT: Kuo, Mei-Chang:
CC APPLICANT: Yeung, Siu-mei H.:
CC APPLICANT: Brauer, Andrew:
CC APPLICANT: Exley, Mark A.:
CC APPLICANT: Powers, Steven P.
CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
CC NUMBER OF SEQUENCES: 261
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
CC STREET: 610 Lincoln St
CC CITY: Waltham
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02154
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/468,940
CC FILING DATE:
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/350,225
CC FILING DATE: December 6, 1994
CC APPLICATION NUMBER: 08/226,248
CC FILING DATE: April 8, 1994
CC APPLICATION NUMBER: 07/338,990
CC FILING DATE: September 11, 1992
CC APPLICATION NUMBER: PC7/US93/00139
CC FILING DATE: January 15, 1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Darlene A. Vanstone
CC REGISTRATION NUMBER: 35,729
CC REFERENCE/DOCKET NUMBER: 025.6 US (IMI-028C22)
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 466-6000
CC TELEFAX: (617) 466-6040
CC INFORMATION FOR SEQ ID NO: 83:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 26 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC FRAGMENT TYPE: internal
CC SEQUENCE 26 AA: 2944 MW: 2965 CN
SQ

CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
CC TITLE OF INVENTION: Japanese Cedar Pollen
CC NUMBER OF SEQUENCES: 201
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
CC STREET: 610 Lincoln St
CC CITY: Waltham
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02154
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/426,248A
CC FILING DATE: April 8, 1994
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/518,490
CC FILING DATE: September 11, 1992
CC APPLICATION NUMBER: PC7/US93/00139
CC FILING DATE: January 15, 1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Darlene A. Vanstone
CC REGISTRATION NUMBER: 35,729
CC REFERENCE/DOCKET NUMBER: 025.5 US (IMI-028C2P)
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 466-6000
CC TELEFAX: (617) 466-6040
CC INFORMATION FOR SEQ ID NO: 83:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 26 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC FRAGMENT TYPE: internal
CC SEQUENCE 26 AA: 2944 MW: 2965 CN
SQ
Query Match 100.0% Score 104.00 Length 26
Best Local Similarity 100.0% Pred. No. 4,81e-05
Matches 157 Conservative 0 Mismatches 0 Indels 0 Gaps 0
DB 1 KSMKVTVAENQFNP 15
Q7 1 KSMKVTVAENQFNP 15
RESULT 14
ID US-08-467-697-86 STANDARD: PRI: 25 AA:
XX xxxxxx
AC
XX
DT
DE
Sequence 86, Application US/08467697
Sequence 86, Application US/08467697
GENERAL INFORMATION:
CC APPLICANT: Griffith, Irwin J.:
CC APPLICANT: Pollock, Joanne:
CC APPLICANT: Bond, Julian F.:
CC APPLICANT: Garman, Richard D.:
CC APPLICANT: Kuo, Mei-Chang:
CC APPLICANT: Yeung, Siu-mei H.:
CC APPLICANT: Brauer, Andrew:
CC APPLICANT: Exley, Mark A.:
CC APPLICANT: Powers, Steven P.
CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
CC NUMBER OF SEQUENCES: 261
CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
 CC STREET: 610 Lincoln St
 CC CITY: Waltham
 CC STATE: MA
 CC COUNTRY: USA
 CC ZIP: 02154
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patent Release #1.3, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/467.697
 CC FILING DATE: June 6, 1995
 CC CLASSIFICATION: 424
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/350,225
 CC FILING DATE: December 6, 1994
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Jane E. Remillard
 CC REGISTRATION NUMBER: 38,872
 CC REFERENCE/DOCKET NUMBER: 025.6 USDA (IMI-028CPD4)
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (617) 227-7400
 CC TELEFAX: (617) 227-5941
 CC INFORMATION FOR SEQ ID NO: 86:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 26 amino acids
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: peptide
 CC FRAGMENT TYPE: internal
 CC SEQUENCE 26 AA: 2944 MW: 2965 CN:
 CC
 CC Query Match 100.0%; Score 106; DB 10; Length 26;
 CC Best Local Similarity 100.0%; Pred. No. 4.81e-05;
 CC Matches 15: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC 1 KSMKVTVAFNQFGPN 15
 CC 1 KSMKVTVAFNQFGPN 15
 CC
 CC RESULT 15
 CC US-08-467-006-83 STANDARD PAT: 26 AA.
 CC xxxxxx
 CC
 CC Sequence 83. Application US/08467005
 CC
 CC Sequence 83. Application US/08467005
 CC GENERAL INFORMATION:
 CC APPLICANT: Griffiths, Irwin J.;
 CC APPLICANT: Pollock, Joanne;
 CC APPLICANT: Bond, Julian F.;
 CC APPLICANT: Garman, Richard D;
 CC APPLICANT: Kuo, Mei-Chang;
 CC APPLICANT: Yeung, Siu-mei H.;
 CC APPLICANT: Brauer, Andrew;
 CC APPLICANT: Exley, Mark A.;
 CC APPLICANT: Powers, Steven P.
 CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
 CC TITLE OF INVENTION: Japanese Cedar Pollen
 CC NUMBER OF SEQUENCES: 261
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESS: Immunologic Pharmaceutical Corporation, Inc.
 CC STREET: 610 Lincoln St
 CC CITY: Waltham
 CC STATE: MA
 CC COUNTRY: USA
 CC ZIP: 02154

CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patent Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/467.906
 CC FILING DATE: June 6, 1995
 CC CLASSIFICATION: 424
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/350,225
 CC FILING DATE: December 6, 1994
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Jane E. Remillard
 CC REGISTRATION NUMBER: 38,872
 CC REFERENCE/DOCKET NUMBER: 025.6 USDA (IMI-028CPD4)
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (617) 227-7400
 CC TELEFAX: (617) 227-5941
 CC INFORMATION FOR SEQ ID NO: 83:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 26 amino acids
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: peptide
 CC FRAGMENT TYPE: internal
 CC SEQUENCE 26 AA: 2944 MW: 2965 CN:
 CC
 CC Query Match 100.0%; Score 106; DB 10; Length 26;
 CC Best Local Similarity 100.0%; Pred. No. 4.81e-05;
 CC Matches 15: Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC 1 KSMKVTVAFNQFGPN 15
 CC 1 KSMKVTVAFNQFGPN 15
 CC
 CC Search completed: Mon Jun 19 16:26:18 2000
 CC Job time : 17 secs.


```

ACCESSIONS      26-Aug-1999
REFERENCE        JC2123: PC2065
AUTHORS         Sano, T.; Komiyama, N.; Shimizu, K.; Kusakabe, T.; Morikubo,
                  K.; Kino, K.
JOURNAL         Biochem. Biophys. Res. Commun. (1994) 199:619-625
TITLE           Cloning and sequencing of cDNA coding for Cry j I, a major
                  allergen of Japanese cedar pollen.
CROSS-REFERENCES MIM:294183234
ACCESSION       JC2123
MOLECULE_TYPE   mRNA
RESIDUES        1-374 **label SON
CROSS-REFERENCES GB:D26544; NID:g493631; PID:d106086; PID:g493632
EXPERIMENTAL_SOURCE pollen
ACCESSION       PC2065
MOLECULE_TYPE   protein
RESIDUES        22-53758-81:219-232:236-259:299-307:346-372 **label SO2
NOTE            the authors described carbohydrate binding site for
                  residue 279
CLASSIFICATION  *superfamily pectate lyase LAt59
KEYWORDS        glycoprotein; pollen
FEATURE         1-21
                  *domain: signal sequence *status: predicted *label: SIGN
                  *product: major allergen Cry j I (clone PCOI-2-2) *status:
                  predicted *label: MATV
158-191,293,354 *binding site carbohydrate (Asn) (covalent) *status:
                  predicted
SUMMARY         #length 374 #molecular-weight 40645 #checksum 2920
                  Query Match 100.0%; Score 106; DB 2; Length 374;
                  Best Local Similarity 100.0%; Pred. No. 1,14e-11;
                  Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DE 232 KSMKVTVAFNQFGPN 245
   1 1111111111111111
QY 1 KSMKVTVAFNQFGPN 15

RESULT 3
ENTRY   S12209 *type complete
TITLE   pectate lyase (EC 4.2.2.2) - tomato
ORGANISM *formal_name Lycopersicon esculentum; *common_name tomato
DATE     19-Mar-1997 *sequence_revision 19-Mar-1997 *text_change
        23-Jul-1999
ACCESSIONS S12209
REFERENCE   S12209
AUTHORS     Roesler, K.A.; Smith, A.G.; Gasser, G.
JOURNAL     Mol. Gen. Genet. (1990) 224:183-192
TITLE       Regulation of a stylar transmission tissue-specific gene in
            wild-type and transgenic tomato and tobacco.
CROSS-REFERENCES MIM:29117185
ACCESSION   S12209
MOLECULE_TYPE mRNA
RESIDUES    1-404 **label BUD
CROSS-REFERENCES GB:X55193; NID:g19161; PID:CAA38979.1; PID:g19162
CLASSIFICATION *superfamily pectate lyase LAt59
KEYWORDS      carbon-oxygen lyase
SUMMARY       #length 404 #molecular-weight 44298 #checksum 4917
                  Query Match 75.5%; Score 80; DB 2; Length 404;
                  Best Local Similarity 76.9%; Pred. No. 3.38e-05;
                  Matches 10; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

DE 262 KMQVTIAFNHFG 274
   1 1111111111111111
QY 1 KSMKVTVAFNQFG 13

RESULT 4
ENTRY   T05556 *type complete
TITLE   pectate lyase (EC 4.2.2.2) F22K18.20 - Arabidopsis thaliana
ALTERNATE_NAMES prote.in F22K18.20

```

```

ORGANISM        *formal_name Arabidopsis thaliana; *common_name model plant;
                  *cross_sequence_revision 23-Apr-1999 *text_change
DATE            23-Jul-1999
ACCESSIONS      T05556
REFERENCE        215419
AUTHORS         Bevan, M.; Medler, H.; Medler, E.; Wambutt, R.; Hohn, M.;
                  Mewes, H.W.; Mayer, K.F.X.; Schueller, C.
SUBMISSION      Submitted to the Protein Sequence Database, February 1999
ACCESSION       T05556
MOLECULE_TYPE   DNA
RESIDUES        1-404 **label BEV
CROSS-REFERENCES EMBL:AL035556
EXPERIMENTAL_SOURCE cultivar Columbia; BA7 clone F22K18
GENETICS
MAP_POSITION    37/2; 267/2
INTRONS         F22K18.20
NOTE            *superfamily pectate lyase LAt59
CLASSIFICATION  *superfamily pectate lyase LAt59
KEYWORDS        carbon-oxygen lyase
SUMMARY         #length 404 #molecular-weight 44495 #checksum 5547
                  Query Match 73.6%; Score 78; DB 2; Length 404;
                  Best Local Similarity 69.2%; Pred. No. 9.92e-05;
                  Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

DE 266 KMQVTIAFNHFG 278
   1 1111111111111111
QY 1 KSMKVTVAFNQFG 13

RESULT 5
ENTRY   T00856 *type complete
TITLE   pectate lyase (EC 4.2.2.2) T20F6.14 - Arabidopsis thaliana
ORGANISM *formal_name Arabidopsis thaliana; *common_name model plant;
                  *cross_sequence_revision 12-Feb-1999 *text_change
DATE     12-Feb-1999 *seq
ACCESSIONS T00856
REFERENCE   214206
AUTHORS     Rounsley, S.D.; ... X.; Ketchum, K.A.; Crosby, M.;
                  Brandon, R.C.; Sykes, S.M.; Kaul, S.; Mason, J.M.;
                  Kertavagay, A.R.; Adams, M.D.; Somerville, C.R.; Venter,
                  J.C.
SUBMISSION  Submitted to the EMBL Data Library, March 1998
DESCRIPTION Arabidopsis thaliana cell suspension; BA7 12-16-1999
ACCESSION   T00856
MOLECULE_TYPE DNA
RESIDUES    1-455 **label MOG
CROSS-REFERENCES EMBL:AC002521; NID:g2947056; PID:AA005150.1;
                  PID:g2947069
EXPERIMENTAL_SOURCE cultivar Columbia
GENETICS
MAP_POSITION    66/2; 295/3; 376/3
INTRONS         T20F6.14
NOTE            *superfamily pectate lyase LAt59
KEYWORDS        carbon-oxygen lyase
SUMMARY         #length 455 #molecular-weight 51257 #checksum 611
                  Query Match 71.7%; Score 76; DB 2; Length 455;
                  Best Local Similarity 69.2%; Pred. No. 2.88e-04;
                  Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

DE 309 KMQVTIAFNHFG 321
   1 1111111111111111
QY 1 KSMKVTVAFNQFG 13

RESULT 6
ENTRY   S26211 *type complete

```

```

TITLE      pectate lyase (EC 4.2.2.2) - common tobacco
FORMAL_NAME Nicotiana tabacum #common_name common tobacco
DATE       12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change
29-Sep-1999
ACCESSIONS S26211; S26212; S21933; S22753; S22754
REFERENCE   S26211; S26212; S21933; S22753; S22754
AUTHORS     Rogers, H.J.; Harvey, A.; Lonsdale, D.M.
JOURNAL     Plant Mol. Biol. (1992) 20:493-502
TITLE       Isolation and characterization of a tobacco gene with
           homology to pectate lyase which is specifically expressed
           during microsporogenesis.
CROSS-REFERENCES EMBL: X67158; NID: g19507; PID: g19508
ACCESSIONS S26211
MOLECULE_TYPE DNA
RESIDUES    1-397 #label ROC
CROSS-REFERENCES EMBL: X67158; NID: g19507; PID: CAA47630.1; PID: g19508
ACCESSION S26212
MOLECULE_TYPE mRNA
RESIDUES    119-155; 157-188; 190-196; 191-199; 200-203; 248-250; 338 #label ROC
CROSS-REFERENCES EMBL: X67159; NID: g19508; PID: CAA47630.1; PID: g19508
REFERENCE   S21933
AUTHORS     Lonsdale, D.M.
SUBMISSION Submitted to the EMBL Data Library, July 1991
ACCESSION S21933
MOLECULE_TYPE DNA
RESIDUES    1-397 #label LON
CROSS-REFERENCES EMBL: X61102; NID: g19508; PID: CAA47630.1; PID: g19508
GENETICS
INTRONS     197/1; 293/2
CLASSIFICATION #superfamily pectate lyase LAT59
KEYWORDS     carbon-oxygen lyase
SUMMARY      #length 397 #molecular-weight 4495; #checksum 1480
Query Match 69.8%; Score 74; DB 2; Length 397;
Best Local Similarity 61.5%; Pred. No. 8,246-04;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
DB 254 KMKVTVAFNFG 266
1 11111111
QY 1 KSMKVTVAFNFG 13
RESULT 7
ENTRY      10524 #type complete
TITLE      pectate lyase (EC 4.2.2.2) - alfalfa
FORMAL_NAME Medicago sativa #common_name alfalfa
DATE       16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change
26-Aug-1999
ACCESSIONS T09524
REFERENCE   T09524
AUTHORS     Wu, Y.; Qiu, X.; Du, S.; Erickson, L.
SUBMISSION Submitted to the EMBL Data Library, November 1995
ACCESSION T09524
STATUS      preliminary; translated from GB/EMBL/DBDJ
MOLECULE_TYPE DNA
RESIDUES    1-450 #label WUY
CROSS-REFERENCES EMBL: J04142; NID: g1171160; PID: g1171161
GENETICS
INTRONS     65/2; 245/3
FUNCTION     catalyzes cleavage of pectate to oligosaccharides
CLASSIFICATION #superfamily pectate lyase LAT59
KEYWORDS     carbon-oxygen lyase
SUMMARY      #length 450 #molecular-weight 50155 #checksum 3091
Query Match 69.8%; Score 74; DB 2; Length 450;
Best Local Similarity 69.2%; Pred. No. 8,246-04;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
DB 309 KMKVTVAFNFG 321
1 11111111

```

```

QY 1 KSMKVTVAFNFG 13
RESULT 8
ENTRY      T07058 #type complete
TITLE      pectate lyase (EC 4.2.2.2) LAT56 - tomato
ALTERNATE_NAMES protein p56
ORGANISM    #formal_name Lycopersicon esculentum #common_name tomato
DATE       30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change
29-Sep-1999
ACCESSIONS T07058; S08383
REFERENCE   T07058; S08383
AUTHORS     Wang, R.A.
SUBMISSION Submitted to the EMBL Data Library, September 1994
ACCESSION T07058
STATUS      translated from GB/EMBL/DBDJ
MOLECULE_TYPE DNA
RESIDUES    1-398 #label WIN
CROSS-REFERENCES EMBL: X15110; NID: g551555; PID: CAA31524.1;
           #experimental_source strain: VP36; author:
           S08383
AUTHORS     Wang, R.A.; Yamauchi, C.; Darabell, S.K.; Driscoll, V.M.;
           McCormick, S.
JOURNAL     Plant Mol. Biol. (1989) 14:17-28
TITLE       Molecular and genetic characterization of two
           pollen-expressed genes that have sequence similarity to
           pectate lyases (the plant pathogen Erwinia
           #cross-references MUID: 91322485
ACCESSION S08383
MOLECULE_TYPE DNA
RESIDUES    1-24; 126-129; 131-132; 133-134; 135-136; 137-138; 139-140; 141-142; 143-156; SNSQCYFIRS-177-349; MVCH-
           354-355; 357 #label WIN
CROSS-REFERENCES EMBL: X15500
GENETICS
MAP_POSITION 194/1; 234/2
INTRONS
CLASSIFICATION #superfamily pectate lyase LAT59
KEYWORDS     carbon-oxygen lyase
SUMMARY      #length 358 #molecular-weight 44563 #checksum 5206
Query Match 57.9%; Score 72; DB 2; Length 358;
Best Local Similarity 53.8%; Pred. No. 2,410-03;
Matches 7; Conservative 6; Mismatches 0; Indels 0; Gaps 0;
DB 455 KMKVTVAFNFG 267
1 11111111
QY 1 KSMKVTVAFNFG 13
RESULT 9
ENTRY      S29612 #type complete
TITLE      pectate lyase (EC 4.2.2.2) - trumpet lily
FORMAL_NAME Lilium longiflorum #common_name trumpet lily
DATE       19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change
29-Sep-1999
ACCESSIONS S29612
REFERENCE   S29612
AUTHORS     Kim, S.; Finkbeiner, D.J.; An, G.
SUBMISSION Submitted to the EMBL Data Library, October 1992
DESCRIPTION Abundance patterns of lily pollen cDNAs: characterization of
           three pollen-preferential cDNA clones.
ACCESSION S29612
STATUS      preliminary
MOLECULE_TYPE mRNA
RESIDUES    1-434 #label KIM
CROSS-REFERENCES EMBL: L217328; NID: g19450; PID: CAA78976.1; PID: g19450;
           #experimental_source cv: Nellie white, mature flower
           #superfamily pectate lyase LAT59
CLASSIFICATION #superfamily pectate lyase LAT59
KEYWORDS     carbon-oxygen lyase
SUMMARY      #length 434 #molecular-weight 48457 #checksum 6432

```

Query Match 67.9% Score 72: DB 2: Length 434:
 Best Local Similarity 81.8% Pred. No. 2.33e-03:
 Matches 9: Conservative 1: Mismatches 1: Indels 0: Gaps 0:

DB 296 MKQVTAIYNHFG 306
 1 1 1 1 1 1 1 1 1 1
 QY 3 MKQVTAIYNHFG 13

RESULT 10
 ENTRY 543335 #type complete
 TITLE pectate lyase (EC 4.2.2.2) - maize
 ALTERNATE_NAMES #formal_name zea mays #common_name maize
 ORGANISM 07-Sep-1994 #sequence_revision 10-Nov-1995 #text_change 29-Sep-1999
 DATE 543335
 ACCESSIONS S43335
 REFERENCE S43334
 #authors Turcich, M.P.; Hamilton, D.A.; Mascarenhas, J.P.
 #journal Plant Mol. Biol. (1999) 23:1061-1065
 #title Isolation and characterization of pollen-specific maize genes with sequence homology to radweed allergens and pectate lyases.
 #cross-references EMBL:U404358
 #accession S43335
 #status nucleic acid sequence not shown
 #molecule_type DNA
 #residues 1-438 #label TUR
 #cross-references EMBL:L26140; NID:9405534; PID:AAAI6476.1; PID:9405535
 CLASSIFICATION #superfamily pectate lyase LAT59
 KEYWORDS carbon-oxygen lyase
 SUMMARY #length 438 #molecular-weight 49138 #checksum 1084

Query Match 67.9% Score 72: DB 2: Length 434:
 Best Local Similarity 81.8% Pred. No. 2.33e-03:
 Matches 9: Conservative 1: Mismatches 1: Indels 0: Gaps 0:

DB 300 MKQVTAIYNHFG 510
 1 1 1 1 1 1 1 1 1 1
 QY 3 MKQVTAIYNHFG 13

RESULT 11
 ENTRY 106728 #type complete
 TITLE pectate lyase (EC 4.2.2.2) F28P10 Arabidopsis thaliana
 ALTERNATE_NAMES Protein F28P10.100
 ORGANISM #formal_name Arabidopsis thaliana #common_name mouse-ear cress
 DATE 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 24-Jul-1999
 ACCESSIONS T05728
 REFERENCE Z15793
 #authors Quetier, F.; Choinsne, N.; Robert, C.; Brettiere, P.; Wincker, P.; Cattolico, C.; Artiguenave, F.; Saurin, W.; Weissenbach, J.; Salanoubat, M.; Mewes, H.W.; Mayer, K.F.X.; Schaefer, C.
 #submission submitted to the Protein Sequence Database, April 1999
 #accession T05728
 #molecule_type DNA
 #residues 1-542 #label CUE
 #cross-references EMBL:AL049655
 #experimental_source cultivar Columbia: BAC clone F28P10

GENETICS
 #map_position 1
 #introns 46/2: 346/3: 413/2: 480/2
 #note F28P10.100
 CLASSIFICATION #superfamily pectate lyase LAT59
 KEYWORDS carbon-oxygen lyase
 SUMMARY #length 542 #molecular-weight 58573 #checksum 1032

Query Match 67.0% Score 71: DB 2: Length 542:
 Best Local Similarity 66.7% Pred. No. 3.89e-03:
 Matches 8: Conservative 3: Mismatches 1: Indels 0: Gaps 0:

DB 280 GMQVTAIYNHFG 291
 1 1 1 1 1 1 1 1 1 1
 QY 2 SMKVTAIYNHFG 13

RESULT 12
 ENTRY 107701 #type complete
 TITLE pectate lyase (EC 4.2.2.2) F17N18.100 Arabidopsis thaliana
 ALTERNATE_NAMES Protein F17N18.100
 ORGANISM #formal_name Arabidopsis thaliana #common_name Arabidopsis thaliana
 DATE 14-May-1999 #sequence_revision 14-May-1999 #text_change 23-Jul-1999
 ACCESSIONS T07701
 REFERENCE Z15098
 #authors Bevan, M.; Pohl, L.; Weizenecker, E.; Bamzait, E.; Mewes, H.W.; Mayer, K.F.X.; Schaefer, C.
 #submission submitted to the Protein Sequence Database, May 1999
 #accession T07701
 #molecule_type DNA
 #residues 1-418 #label BEV
 #cross-references EMBL:AL049751
 #experimental_source cultivar Columbia: BAC clone F17N18

GENETICS
 #map_position 4
 #introns 26/2: 49/2: 264/3: 289/2: 345/3
 #note F17N18.100
 CLASSIFICATION #superfamily pectate lyase LAT59
 KEYWORDS carbon-oxygen lyase
 SUMMARY #length 418 #molecular-weight 46251 #checksum 599

Query Match 66.0% Score 70: DB 2: Length 418:
 Best Local Similarity 61.5% Pred. No. 6.45e-03:
 Matches 8: Conservative 3: Mismatches 2: Indels 0: Gaps 0:

DB 278 KMQVTAIYNHFG 290
 1 1 1 1 1 1 1 1 1 1
 QY 1 KMKVTAIYNHFG 13

RESULT 13
 ENTRY T05240 #type complete
 TITLE pectate lyase (EC 4.2.2.2) Arabidopsis thaliana
 ALTERNATE_NAMES Protein F18A5.100
 ORGANISM #formal_name Arabidopsis thaliana #common_name mouse-ear cress
 DATE 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 23-Jul-1999
 ACCESSIONS T05240
 REFERENCE Z15405
 #authors Bevan, M.; Weber, N.; Grueninger, D.; Schmidhafer, T.; Bacroff, I.; Mewes, H.W.; Mayer, K.F.X.; Schaefer, C.
 #submission submitted to the Protein Sequence Database, February 1999
 #accession T05240
 #molecule_type DNA
 #residues 1-374 #label BEV
 #cross-references EMBL:AL035528
 #experimental_source cultivar Columbia: BAC clone F18A5

GENETICS
 #map_position 4
 #introns positions not resolved
 #note F18A5.100
 CLASSIFICATION #superfamily pectate lyase LAT59
 KEYWORDS carbon-oxygen lyase
 SUMMARY #length 374 #checksum 1983

Query Match 64.2% Score 68: DB 2: Length 374:
 Best Local Similarity 61.5% Pred. No. 1.78e-02:
 Matches 8: Conservative 3: Mismatches 2: Indels 0: Gaps 0:

DB 234 KLMQVTAIYNHFG 246

CC AXB A 1/AXB A 11/CRY J 1 SUSFAMILY

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL: L26544; BAA05542.1; 1

CC EMBL: S25545; BAA05543.1; 1

CC EMBL: D34639; BAA07020.1; 1

CC EMBL: A44773; A44773.1

CC PFAM: PF00544; pec_lyase.1

CC PRINTS: PR00807; AMBALLERGEN

CC Allergens: Glycoprotein; Multigene family; Signal.

CC SIGNAL: 21

CC CHAIN: 22 374

CC VARIANT: 12 12

CC VARIANT: 143 143

CC VARIANT: 222 222

CC VARIANT: 221 221

CC VARIANT: 358 358

CC VARIANT: 361 361

CC VARIANT: 158 158

CC CARBOHYD: 191 191

CC CARBOHYD: 191 191

CC CARBOHYD: 293 293

CC SEQUENCE: 374 AA; 45645 MW; 74AB29950248F56F CRC64;

CC

CC Query Match 100.0% Score 166 DB 11 Length 374

CC Best Local Similarity 100.0% Pred. No. 1.8 Gaps 0

CC Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC

CC 232 KSMKVTVAFNQSPN 246

CC : KSMKVTVAFNQSPN 15

CC

CC RESULT 3

CC HL 9612_LYCES STANDARD: PRT: 404 AA

CC P24356:

CC 01-MAR-1992 (Rel. 21; Created)

CC 01-MAR-1992 (Rel. 21; Last sequence update)

CC 15-JUL-1995 (Rel. 38; Last annotation update)

CC STYLE DEVELOPMENT-SPECIFIC PROTEIN 9612 PRECURSORS R

CC 9612

CC Lycopersicon esculentum (Tomato)

CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

CC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;

CC core eudicots; Asteridae; euasterids I; Solanales; Solanaceae;

CC Solanum

CC [1]

CC SEQUENCE FROM N.A.

CC STRAIN-CV: VF35; TISSUE-P: STIL;

CC MEDLINE: 91117185

CC Buedeler K.A., Smith A.G., Gasser C.S.;

CC Regulation of a stylar transmitting tissue-specific gene in

CC wild-type and transgenic tomato and tobacco.*;

CC Mol. Gen. Genet. 224:183-192(1990).

CC FUNCTION: MAY HAVE A ROLE IN THE DEVELOPMENT OF THE TRANSMITTING

CC TISSUE OF THE STYLE AND/OR IN THE EVENTS RELATED TO POLLINATION

CC SUCH AS SOME ASPECT IN THE FACILITATION OF COMPATIBLE POLLEN TUBE

CC GROWTH.

CC

CC -1- SUBCELLULAR LOCATION: SECRETED.

CC -1- TISSUE SPECIFICITY: PREDOMINANTLY FOUND IN THE PISTIL WHERE IT IS

CC FOUND IN THE OUTER FIVE LAYERS OF THE STRANDS OF TRANSMITTING

CC TISSUE WITHIN THE UPPER TWO-THIRDS OF THE STYLE. FOUND AT MUCH

CC LOWER LEVELS IN THE ANTHEIRS AND VEGETATIVE ORGANS

CC -1- DEVELOPMENTAL STAGE: MAXIMUM LEVELS ARE FOUND DURING ANTHESIS.

CC -1- SIMILARITY: 54% IDENTICAL TO TOMATO PROTEINS P59 (AC P15722)

CC AND P56 (AC P15721).

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL: X55193; CAA38379.1; 1

CC EMBL: S22259; S12209.1

CC PFAM: PF00544; pec_lyase.1

CC PRINTS: PR00807; AMBALLERGEN

CC SIGNAL: 1 20

CC CHAIN: 21 404

CC CARBOHYD: 37 37

CC CARBOHYD: 191 191

CC SEQUENCE: 404 AA; 44298 MW; F46ED49H:28D8675 CRC64;

CC

CC Query Match 75.5% Score 80 DB 1 Length 404

CC Best Local Similarity 76.9% Pred. No. 3.02e-06

CC Matches 10; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

CC

CC Db 262 KGMQVTVAFNQSPN 274

CC : KSMKVTVAFNQSPN 13

CC

CC RESULT 3

CC LD PEL-TOBAC STANDARD: PRT: 497 AA

CC P43972:

CC 01-FEB-1995 (Rel. 31; Created)

CC 01-FEB-1995 (Rel. 31; Last sequence update)

CC 15-DEC-1998 (Rel. 37; Last annotation update)

CC PECTATE LYASE PRECURSOR (EC 4.2.2.2)

CC DE

CC Nicotiana tabacum (Common tobacco)

CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

CC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;

CC core eudicots; Asteridae; euasterids I; Solanales; Solanaceae;

CC Nicotiana

CC [1]

CC SEQUENCE FROM N.A.

CC STRAIN-CV: SAMSUN; TISSUE-P: N;

CC MEDLINE: 93043039

CC Rogers H.C., Harvey A.L., Rogers H.C., Rogers H.C.;

CC Isolation and characterization of a cDNA encoding a putative

CC pectate lyase which is specifically expressed during

CC microsporogenesis.

CC Plant Mol. Biol. 20:493-504(1992)

CC -1- CATALYTIC ACTIVITY: ELIMINATIVE CLEAVAGE OF PECTATE TO GIVE

CC GLUCOSACCHARIDES WITH 4-DEOXY-ALPHA-D-XANN-4-ENURONOSYL GROUPS

CC AT THEIR NON-REDUCING ENDS.

CC -1- DEVELOPMENTAL STAGE: MAXIMALLY EXPRESSED LATE IN POLEN

CC DEVELOPMENT.

CC

CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see <http://www.isb-sib.ch/announcement/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL: X57158; CAA47630.1; 1

CC EMBL: X57159; CAA47631.1; 1

CC EMBL: X61102; CAA434.4.1; 1

CC PIR: S26211; S26211

CC PFAM: PF00544; pec_lyase.1

CC PRINTS: PR00807; AMBALLERGEN

CC Lyase: Signal.

CC -1- SIMILARITY: 1 25

CC SIGNAL

CC CHAIN 26 397

CC PECTATE LYASE.

FT ACT_SITE 272 272 POTENTIAL.
 FT CARBOHYD 134 134 POTENTIAL.
 FT CARBOHYD 227 227 POTENTIAL.
 FT CONFLICT 156 156 S -> C (IN MRNA).
 FT CONFLICT 189 190 GS -> SG (IN MRNA).
 FT CONFLICT 200 200 S -> D (IN MRNA).
 FT CONFLICT 202 202 H -> R (IN MRNA).
 FT CONFLICT 249 249 H -> N (IN MRNA).
 SQ SEQUENCE 397 AA: 44351 MW: EFC482CE5DA7643F CRC64:

Query Match 59.8%; Score 74; DB 1; Length 397;
 Best Local Similarity 61.5%; Pred. No. 1,05e-04;
 Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

DB 254 KMKITLAYNHFG 266
 QY 1 KSMKVTVAFNQFG 13

RESULT 4
 ID PE55_LYCES STANDARD: PRT: 398 AA.

AC P15721.
 DT 01-APR-1990 (Rel. 14, Created).
 DT 01-FEB-1996 (Rel. 33, Last sequence update).
 DT 15-JUL-1999 (Rel. 38, Last annotation update).
 DE PROBABLE PECTATE LYASE P56 PRECURSOR (EC 4.2.2.2).
 GN LAT56.
 OS Lycopersicon esculentum (Tomato).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 CC core eudicots; Asteridae; easterids I; Solanales; Solanaceae;
 CC Solanum.
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-CV, VF16; TISSUE=ANTHER;
 RX MEDLINE: 91322485
 RA Wang R.A., Yamaguchi J., Larabell S.K., Frisn V.M., McCormick S.;
 RT "Molecular and genetic characterization of two pollen-expressed genes
 R1 that have sequence similarity to pectate lyases of the plant pathogen
 R2 Erwinia";
 R3 Plant Mol. Biol. 14:17-28(1990).
 RN [2]

RP REVISIONS.
 RC STRAIN-CV, VF16; TISSUE=ANTHER;
 RA Wang R.A.;
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
 CC FUNCTION: MIGHT BE NEEDED DURING POLLEN DEVELOPMENT. AND TUBE
 CC GROWTH.

CC CATALYTIC ACTIVITY: ELIMINATIVE CLEAVAGE OF PECTATE TO GIVE
 CC OLIGOSACCHARIDES WITH 4-DEOXY-ALPHA-D-MANN-4-ENURONOSYL GROUPS
 CC AT THEIR NON-REDUCING ENDS.
 CC TISSUE SPECIFICITY: EXPRESSED IN ANTHERS AND POLLEN.
 CC SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DB EMBL: X15500; CAA33524.1;
 DR PIR: S08383; S08383.
 DR PFAM: PF00544; pec_lyase.1.
 DR PRINTS: PR00807; AMBALLERGEN.
 KW Lyase: Multigene family; Signal;
 FT SIGNAL 1 27 OR 22 (POTENTIAL).
 FT CHAIN 28 398 PROBABLE PECTATE LYASE P56.
 FT ACT_SITE 273 273 POTENTIAL.
 FT CARBOHYD 135 135 POTENTIAL.
 FT CARBOHYD 228 228 POTENTIAL.
 SQ SEQUENCE 398 AA: 44563 MW: 8D676250BD8BC7C8 CRC64:

Query Match 67.9%; Score 72; DB 1; Length 398;
 Best Local Similarity 53.8%; Pred. No. 3.34e-04;
 Matches 7; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

DB 255 RGMKITLAYNHFG 267
 QY 1 KSMKVTVAFNQFG 13

RESULT 5
 ID PE5_LILLO STANDARD: PRT: 434 AA.

AC P40973.
 DT 01-FEB-1995 (Rel. 31, Created).
 DT 01-FEB-1995 (Rel. 31, Last sequence update).
 DT 15-DEC-1998 (Rel. 37, Last annotation update).
 DE PECTATE LYASE PRECURSOR (EC 4.2.2.2).
 OS Lilium longiflorum (Trumpet lily).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida; Liliales;
 CC Liliaceae; Lilium.
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN-CV, NELLIE WHITE; TISSUE=POLLEN.
 RA Kim S.R., Finkel D.J., An G.;
 RL Submitted (JUN-1993) to the EMBL/GenBank/DBJ databases.
 CC CATALYTIC ACTIVITY: ELIMINATIVE CLEAVAGE OF PECTATE TO GIVE
 CC OLIGOSACCHARIDES WITH 4-DEOXY-ALPHA-D-MANN-4-ENURONOSYL GROUPS
 CC AT THEIR NON-REDUCING ENDS.

CC SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DB EMBL: Z17328; CAA78976.1;
 DR EMBL: L18911; AAA33398.1;
 DR PIR: S29612; S29612.
 DR PFAM: PF00544; pec_lyase.1.
 DR PRINTS: PR00807; AMBALLERGEN.
 KW Lyase: Signal;
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 23 434 PECTATE LYASE.
 FT ACT_SITE 312 312 POTENTIAL.
 FT CARBOHYD 68 68 POTENTIAL.
 FT CARBOHYD 97 97 POTENTIAL.
 SQ SEQUENCE 434 AA: 48457 MW: C1F3E30AD2B5D064 CRC64:

Query Match 67.9%; Score 72; DB 1; Length 434;
 Best Local Similarity 61.8%; Pred. No. 3.34e-04;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

DB 296 MKVTVAFNHFG 306
 QY 3 KMKVTVAFNQFG 13

RESULT 6
 ID PE59_LYCES STANDARD: PRT: 449 AA.

AC P15722.
 DT 01-APR-1990 (Rel. 14, Created).
 DT 01-APR-1990 (Rel. 14, Last sequence update).
 DT 15-JUL-1999 (Rel. 38, Last annotation update).
 DE PROBABLE PECTATE LYASE P59 PRECURSOR (EC 4.2.2.2).
 GN LAT59.
 OS Lycopersicon esculentum (Tomato).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 CC core eudicots; Asteridae; easterids I; Solanales; Solanaceae;

RN Solanum.
 RP [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN:CV, VF35; TISSUE:ANTHER;
 RX MEDLINE: 9322485.
 RA Wang R.A., Yabuchi S., Liorbelli S.K., Ursin V.M., McCormick S.,
 RI Molecular and genetic characterization of two pollen-expressed genes
 RI that have sequence similarity to pectate lyases of the plant pathogen
 RI Erwinia.
 RI Plant Mol. Biol. 14:17-28(1990).
 CC [1] FUNCTION: MIGHT BE NEEDED DURING POLLEN DEVELOPMENT AND TUBE
 CC GROWTH.
 CC [1] CATALYTIC ACTIVITY: ELIMINATIVE CLEAVAGE OF PECTATE TO GIVE
 CC OLIGOSACCHARIDES WITH 4-DEOXY-ALPHA-D-MANN-4-ENURONSYL GROUPS
 CC AS THEIR NON-REDUCING ENDS.
 CC [1] TISSUE SPECIFICITY: EXPRESSED IN ANTHERS AND POLLEN.
 CC [1] SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: X15499; CAA13523.1;
 CC PIR: S27098; S27098.
 CC PIR: P00344; P00344.
 CC PRINTS: P00307; AMBALLERGEN.
 CC Lyase: P00307; Signal.
 CC SIGNAL 22 POTENTIAL.
 CC CHAIN 21 449 PROBABLE PECTATE LYASE P59.
 CC ACT SITE 325 325 POTENTIAL.
 CC CARBOXYD 56 56 POTENTIAL.
 CC CARBOXYD 80 80 POTENTIAL.
 CC CARBOXYD 81 81 POTENTIAL.
 CC SEQUENCE 449 AA: 50893 MW: 1782AA13F:735C3C C664;
 Query Match 60.4% Score 64; DB 1; Length 449;
 Best Local Similarity 61.6% Pred. No. 2,826-62.
 Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 DB 309 MOLLUSCA:VF35
 CC [1]
 CC 3 MKVIVAFN0FG 13
 RESULT 7
 AC G3PA GRAVE STANDARD: PIR: 416 AA
 AC P30724;
 DI 01-APR-1993 (Rel. 25, Created)
 DI 01-APR-1993 (Rel. 25, Last sequence update)
 DI 15-JUL-1999 (Rel. 38, Last annotation update)
 DE EUKARYOTIC PEPTIDE CHAIN RELEASE FACTOR SUBUNIT 1 (ERFL) (UNIPROT)
 DE GLYCERALDEHYDE 3-PHOSPHATE DEHYDROGENASE, CHLOROPLAST PRECURSOR
 DE (EC 1.2.1.12).
 GN GAP.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophytes;
 OC Euphyllophytes; Spermatophytes; Magnoliopsida; Eudicotyledons;
 OC Core eud. cots.; Rosidae; Eustoids 1; Brassicales; Brassicaceae;
 OC Arabidopsis.
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN:CV, COLUMBIA;
 RA Brown C.M., Quigley F.R., Miller W.A.;
 RI Three eukaryotic release factor one (erfl) homologs from Arabidopsis
 RI thaliana Columbia.
 RI Plant Gene Register PGR95-123.
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN:CV, COLUMBIA;
 RX MEDLINE: 93286699.
 RA Quigley F., Dao P., Cottet A., Maché R.;
 RI "Sequence analysis of an 81 kb contig from Arabidopsis thaliana

RI glyceroldehyde-3-phosphate dehydrogenase from the marine red alga
 RI Gracilaria verrucosa.
 CC Curr. Genet. 26:79-86(1994).
 CC [1] CATALYTIC ACTIVITY: D-GLYCERALDEHYDE 3-PHOSPHATE + NADPH + H₂O
 CC + NAD(+) -> 1,3-BIPHOSPHATETETRAE + NADH.
 CC [1] PATHWAY: CALVIN CYCLE.
 CC [1] SUBUNIT: HOMOTETRAMER.
 CC [1] SUBCELLULAR LOCATION: CHLOROPLAST.
 CC [1] SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
 CC DEHYDROGENASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: Z15102; CAA7881.1;
 CC EMBL: L22011; AAA33355.1;
 CC PIR: S25596; S25596.
 CC PIR: S45484; S45484.
 CC HSP: P50362; INH.
 CC MENDEL: 560; GRAVE:GAP.1
 CC PRINTS: P00078; G3PDHGNASE.
 CC PRSITE: P00071; GAPDH: 1.
 CC PFAM: P00044; gpdh: 1.
 CC Calvin cycle; Oxidoreductase; NAD; Transit peptide; Chloroplast.
 KW Multigene family.
 FT TRANSIT 1 78
 FT CHAIN 79 416
 FT BINDING 233 233
 FT ACT SITE 260 260
 FT ACT SITE 260 260
 FT SEQUENCE 416 AA: 44337 MW: 13594.7E4230B8 C664;
 Query Match 55.7% Score 59; DB 1; Length 416;
 Best Local Similarity 75.0% Pred. No. 3,830-01;
 Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 DB 75 MKVIVAFN0FG 89
 CC [1]
 CC 2 MKVIVAFN0FG 13
 RESULT 8
 AC EFL1 ARABID TH STANDARD: PIR: 416 AA
 AC P35614;
 DI 01-JUN-1994 (Rel. 29, Created)
 DI 01-JUN-1994 (Rel. 29, Last sequence update)
 DI 15-JUL-1999 (Rel. 38, Last annotation update)
 DE EUKARYOTIC PEPTIDE CHAIN RELEASE FACTOR SUBUNIT 1 (ERFL) (UNIPROT)
 DE SUPPRESSOR PROTEIN 1 HOMOLOG; (SUP1 HOMOLOG).
 GN EFL1-3.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophytes;
 OC Euphyllophytes; Spermatophytes; Magnoliopsida; Eudicotyledons;
 OC Core eud. cots.; Rosidae; Eustoids 1; Brassicales; Brassicaceae;
 OC Arabidopsis.
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN:CV, COLUMBIA;
 RA Brown C.M., Quigley F.R., Miller W.A.;
 RI Three eukaryotic release factor one (erfl) homologs from Arabidopsis
 RI thaliana Columbia.
 RI Plant Gene Register PGR95-123.
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN:CV, COLUMBIA;
 RX MEDLINE: 93286699.
 RA Quigley F., Dao P., Cottet A., Maché R.;
 RI "Sequence analysis of an 81 kb contig from Arabidopsis thaliana

```

FT BINDING 152 152 GLYCERALDEHYDE 3-PHOSPHATE
ACT_SIZE 179 179 ACTIVATES THIOE GROUP DURING CATALYSIS
SQ SEQUENCE 340 AA: 37476 MW: DQDDHBC75633E6 CRC64:

Query Match 53.8% Score 57 DB 1: Length 410
Best Local Similarity 72.7% Pred. No. 1.04e+00:
Matches 8: Conservative 1: Mismatches 2: Indels 3: Gaps 0:

DB 1 MKVKAINGFG 11
1111111111
QY 3 MKVVAINGFG 13

RESULT 10
ID MP:3 AMBAR STANDARD: PRT: 397 AA.
AC P2751:
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE POLLEN ALLERGEN AMB A 1.3 PRECURSOR (ANTIGEN E) (ANTIGEN AMB A 1)
OS Ambrosia artemisiifolia (Short ragweed)
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
OC Euphyllophytes: Spermatophyta: Magnoliophyta: eudicotyledons:
OC core eudicots: Asteridae: euasterids II: Asterales: Asteraceae:
OC Asteroideae: Heliantheae: Ambrosia.
RN [1]
RC SEQUENCE FROM N.A.
RP TISSUE=POLLEN:
RX MEDLINE: 91093235
RA Rafnar T., Griffith I.J., Kuo M.-C., Bond J.F., Rogers H.L.,
RA Klapper D.G.:
RT "Cloning of Amb A 1 (antigen E), the major allergen family of short
RT ragweed pollen."
RL J. Biol. Chem. 265:1229-1235(1991).
RN [2]
RC SEQUENCE FROM N.A., AND VANTS.
RP TISSUE=POLLEN:
RX MEDLINE: 92234570.
RA Griffith I.J., Pollock J., Klapper D.G., Rogers B.L., Nault A.K.:
RA "Sequence polymorphism of Amb A 1 and Amb A 11, the major allergens
RA in Ambrosia artemisiifolia (short ragweed)."
RL Int. Arch. Allergy Appl. Immunol. 96:296-304(1991).
RN [3]
RC TISSUE=MONOMER.
CC [1] TISSUE SPECIFICITY: POLLEN AND FLOWERS
CC [2] DISEASE: ONE OF THE MAJOR ALLERGENS OF THE RAGWEED POLLEN.
CC [3] SIMILARITY: BELONGS TO THE GLYCERALDEHYDE 3-PHOSPHATE
CC AMB A 1/AMB A 11/CRY J 1 SUBFAMILY
CC THIS SWISS-PROT entry is Copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch)
CC
CC EMBL: M62961: AAA32668.1:
CC EMBL: M80580: AAA32669.1: ALT_SEQ.
CC PIR: C39099: C39099.
CC PFAM: PF00544: pec1yase. 1.
CC PRINTS: PR08007: AMBALLERGEN.
CC Antigen: Allergen: Signal: Multigene family: Polymorphism.
FT SIGNAL 1 25 POTENTIAL.
FT CHAIN 26 397 POLLEN ALLERGEN AMB A 1.3.
FT MOD_RES 926 726 BLOCKED.
FT VARIANT 48 48 L -> Y.
SQ SEQUENCE 397 AA: 42928 MW: C8D841257590D50A CRC64:

Query Match 53.8% Score 57 DB 1: Length 397;
Best Local Similarity 50.0% Pred. No. 1.04e+00:
Matches 9: Conservative 1: Mismatches 5: Indels 0: Gaps 0:

DB 256 KGMATVAENFTDN 270

```


RP SEQUENCE FROM N.A.
 PA MEDLINE: 87031573.
 PA Paces V., Vitek C., Urbanek P.:
 PA "nucleotide sequence of the late region of *Bacillus subtilis* phage
 PI PZA, a close relative of phi 29.";
 RI Gene 44:107-114(1986).
 CC -1- FUNCTION: HELPS TO RELEASE THE MATURE PHAGE PARTICLES FROM THE
 CC CELL WALL BY BREAKING DOWN THE PEPTIDOGLYCAN.
 CC -1- CATALYTIC ACTIVITY: HYDROLYSES OF THE 1,4-BETA-LINKAGES BETWEEN
 CC N-ACETYL-D-GLUCOSAMINE AND N-ACETYL-MURAMIC ACID IN PEPTIDOGLYCAN
 CC HETEROPOLYMERS OF THE PROKARYOTES CELL WALLS.
 CC -1- SIMILARITY: BELONGS TO FAMILY 24 OF GLYCOSYL HYDROLASES.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: M18133; AAA88432.1;
 DR PIR: A26215; WMBP15.
 DR HSP: PCG720; 205L.
 DR PFAM: PF00959; Phage_lysozyme; 1.
 DR PFAM: PF01476; PG_binding_2; 2.
 KW Late Protein; Hydrolase; Glycosidase; Bacteriolytic enzyme.
 FT ACT SITE 15 PROTON DONOR (BY SIMILARITY).
 FT VARIANTS 15
 FT VARIANTS 265 AA: 28052 MW: 1457505334637169 CRC64:
 SQ SEQUENCE 265 AA: 28052 MW: 1457505334637169 CRC64:
 Query Match 51.9%; Score 55; DB 1; Length 258;
 Best Local Similarity 50.0%; Pred. No. 2.7; -00;
 Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
 DB 70 KALKVSTGNGFDA 83
 QY 1 KSMKVTVAFNQGP 14
 RESULT 14
 ID WP12 AVBAR STANDARD; PRT: 398 AA.
 AC P27750:
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 LE POLLEN ALLERGEN AMB A 1.2 PRECURSOR (ANTIGEN E) (ANTIGEN AMB A 1).
 CS Ambrosia artemisiifolia (short ragweed).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta.
 CC eupyrophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 CC core eudicots; Asterales; Asterids 1; Asterales; Asteraceae;
 CC Asteroideae; Heliantheae; Ambrosia.
 CC [1]
 RN SEQUENCE FROM N.A.
 RP TISSUE-POLLEN;
 RX MEDLINE: 91093235.
 RA Rafnar T., Griffith I.J., Xiao M.-C., Bond J.F., Rogers B.L.,
 RA Klapper D.G.:
 RI "Cloning of Amb A 1 (antigen E), the major allergen family of short
 RI ragweed pollen.";
 RJ J. Biol. Chem. 266:1229-1236(1991).
 RN [2]
 RP SEQUENCE FROM N.A., AND VARIANTS.
 RC TISSUE-POLLEN;
 RX MEDLINE: 92234570.
 RA Griffith I.J., Pollock J., Klapper D.G., Rogers B.L., Nault A.K.:
 RI "Sequence polymorphism of Amb A 1 and Amb A 11, the major allergens
 RI in *Ambrosia artemisiifolia* (short ragweed).";
 RJ Int. Arch. Allergy Appl. Immunol. 96:296-304(1991).
 CC -1- SUBUNIT: MONOMER.
 CC -1- TISSUE SPECIFICITY: POLLEN AND FLOWERS.
 CC -1- DISEASE: ONE OF THE MAJOR ALLERGENS OF THE RAGWEED POLLEN.
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC AMB A 1/AMB A 11/CRY J 1 SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: M52981; AAA32566.1;
 DR EMBL: M80559; AAA32667.1;
 DR PIR: B39099; B39099;
 DR PFAM: PF00544; pect_lyase_1;
 DR PRIN-S: PR08037; AMBAL-EMGEN
 KW Antigen; Allergen; Signal; M; gene family: Polymorphism.
 FT SIGNAL 1 25
 FT CHAIN 26 398 1 GEN ALLERGEN AMB A 1.2.
 FT MOD_RES 226 226 91 XED.
 FT VARIANT 345 345 4 -> K.
 FT VARIANT 381 381 2 -> .
 SQ SEQUENCE 398 AA: 43664 MW: 420076429474150 CRC64:
 Query Match 51.9%; Score 55; DB 1; Length 398;
 Best Local Similarity 66.7%; Pred. No. 2.77e-00;
 Matches 8; Conservative 1; Mismatches 3; Indels 0;
 DB 257 KGLATVAFNMF 268
 QY 1 KSMKVTVAFNQF 12
 RESULT 15
 ID YAMB-SCHPO STANDARD; PRT: 486 AA.
 AC Q10063:
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 DE HYPOTHETICAL 53.9 KD PROTEIN C1F5.36C IN CHROMOSOME 1.
 GN SPAC1F5.08C.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetales;
 OC Schizosaccharomycetaceae; Schizosaccharomycetes.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-972;
 RA Gentile S., O'Neil G., Church M., Fattol, H.G., Randerfer M.A.,
 RA Walsh S.V.:
 RL Submitted (Nov-1995) to the EMBL/GenBank/DBJ databases
 CC -1- SIMILARITY: SOME TO YEAST NUC
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: Z68136; CAA92236.1;
 KW Hypothetical protein; Transmembrane; Glycoprotein.
 FT TRANSMEM 12 32 POTENTIAL.
 FT TRANSMEM 465 485 POTENTIAL.
 FT CARBOHYD 33 33 POTENTIAL.
 FT CARBOHYD 49 49 POTENTIAL.
 FT CARBOHYD 59 59 POTENTIAL.
 FT CARBOHYD 82 82 POTENTIAL.
 FT CARBOHYD 93 93 POTENTIAL.
 FT CARBOHYD 160 160 POTENTIAL.
 FT CARBOHYD 181 181 POTENTIAL.
 FT CARBOHYD 185 185 POTENTIAL.
 FT CARBOHYD 201 201 POTENTIAL.
 FT CARBOHYD 255 255 POTENTIAL.
 FT CARBOHYD 307 307 POTENTIAL.

F1 CARBOHYD 330 330 POTENTIAL.
F1 CARBOHYD 341 341 POTENTIAL.
F1 CARBOHYD 353 353 POTENTIAL.
F1 CARBOHYD 379 379 POTENTIAL.
F1 CARBOHYD 445 445 POTENTIAL.
SL SEQUENCE 486 AA: 53888 MW: 182047356008958 CR564:

Query Match 51.9% Score 55: DB 1: Length 486:
Best Local Similarity 61.6% Pred. No: 2.77e+00:
Matches 7: Conservative 3: Mismatches 1: Indels 0: Gaps 0:

DT 118 VSSAFSCUAPN 128
11111111
QY 5 VIVAFNOFGN 15

Search completed: Mon Jun 19 16:09:08 2000
Total time: 16 secs.

M E R E L

(1M)

Release 3.1A John F. Collins, Biocon using Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Mrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Jun 19 16:09:25 2000: MasPar time 8.84 Seconds
117,642 Million cell updates/sec

Tabular output not generated.

Title: >US-09-142-524A-9
Description: (1-15) from US09142524A pep
Perfect Score: 106
Sequence: 1 KSMKVIVAFNQFGPN 15

Scoring table: PAM 150
Gap 15

Searched: 225878 seqs, 5933422 residues

Post-processing: Minimum Match 08
Listing first 45 summaries

Database: sprenb112
1:sp_archer 2:sp_bacteria 3:sp_fungi 4:sp_human
5:sp_invertebrate 6:sp_mammal 7:sp_mic 8:sp_oranelle
9:sp_plant 10:sp_plant 11:sp_rodet 12:sp_unclassified
13:sp_vertebrate 14:sp_virus

Statistics: Mean 26.072: Variance 31.116: scale 0.838

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description | Pred. No. |
|------------|-------|-------------|--------|-------|-------------------------|-----------|
| 1 | 106 | 100.0 | 367 | 10 | POLLEN MAJOR ALLERGEN | 1.02e-12 |
| 2 | 106 | 100.0 | 375 | 10 | CHAO1 PRECURSOR | 1.02e-12 |
| 3 | 87 | 82.1 | 394 | 10 | PECTATE LYASE LIKE PRO | 1.46e-07 |
| 4 | 87 | 82.1 | 394 | 10 | PECTATE LYASE LIKE PRO | 1.46e-07 |
| 5 | 86 | 81.1 | 390 | 10 | F12F1.22 PROTEIN | 2.67e-07 |
| 6 | 82 | 77.4 | 398 | 10 | PECTATE LYASE (EC 4.2) | 2.87e-06 |
| 7 | 78 | 73.6 | 226 | 10 | PUTATIVE PECTATE LYASE | 2.96e-05 |
| 8 | 78 | 73.6 | 401 | 10 | PECTATE LYASE PRECURSOR | 2.96e-05 |
| 9 | 76 | 71.7 | 455 | 10 | PUTATIVE PECTATE LYASE | 9.30e-05 |
| 10 | 76 | 71.7 | 459 | 10 | PUTATIVE PECTATE LYASE | 9.30e-05 |
| 11 | 75 | 70.9 | 431 | 10 | TIG11.7 PROTEIN | 1.64e-04 |
| 12 | 74 | 69.8 | 263 | 10 | PECTATE LYASE (EC 4.2) | 2.88e-04 |
| 13 | 74 | 69.8 | 450 | 10 | PECTATE LYASE HOMOLOG | 2.88e-04 |
| 14 | 72 | 67.9 | 438 | 10 | POLLEN SPECIFIC PECTAT | 8.82e-04 |
| 15 | 71 | 67.0 | 405 | 10 | PECTATE LYASE | 1.53e-03 |
| 16 | 68 | 64.2 | 227 | 10 | PUTATIVE PECTATE LYASE | 7.88e-03 |
| 17 | 64 | 60.4 | 171 | 10 | PUTATIVE PECTATE LYASE | 6.57e-02 |
| 18 | 64 | 60.4 | 432 | 10 | PUTATIVE ZINC-FINGER P | 6.57e-02 |
| 19 | 64 | 60.4 | 461 | 20 | T15F16.7 PROTEIN | 6.57e-02 |
| 20 | 62 | 58.5 | 210 | 10 | F13F21.9 PROTEIN | 1.85e-01 |

| | | | | | | | |
|----|----|------|------|----|--------|------------------------|----------|
| 21 | 58 | 54.7 | 554 | 5 | OS1966 | F26D11.10 PROTEIN | 1.47e-00 |
| 22 | 57 | 53.8 | 414 | 8 | Q37174 | GLYCERALDEHYDE-3-PH-SP | 2.23e-00 |
| 23 | 57 | 53.8 | 436 | 10 | Q49097 | EUKARYOTIC RELEASE FAC | 2.23e-00 |
| 24 | 57 | 53.8 | 571 | 2 | P77318 | HYPOTHETICAL 54.1 KD P | 2.23e-00 |
| 25 | 57 | 53.8 | 2823 | 5 | Q45614 | T22A3.8 PROTEIN (FRAGM | 2.23e-00 |
| 26 | 55 | 51.9 | 350 | 2 | Q32073 | MULTIDRUG RESISTANCE P | 5.80e-00 |
| 27 | 55 | 51.9 | 808 | 10 | Q23052 | SIMILARITY TO RATIUS C | 5.80e-00 |
| 28 | 54 | 50.9 | 286 | 10 | Q20221 | PUTATIVE ARGININE/SER | 4.26e-00 |
| 29 | 54 | 50.9 | 355 | 8 | Q37453 | NADH DEHYDROGENASE SUB | 9.24e-00 |
| 30 | 54 | 50.9 | 355 | 8 | Q37446 | NADH DEHYDROGENASE SUB | 9.24e-00 |
| 31 | 54 | 50.9 | 883 | 2 | Q53136 | TRANSEFERING BINDING P | 9.24e-00 |
| 32 | 54 | 50.9 | 1434 | 5 | Q62239 | F423.8 PROTEIN | 9.24e-00 |
| 33 | 53 | 50.0 | 216 | 14 | P85112 | ENVELOPE GLYCOPROTEIN | 1.47e-01 |
| 34 | 53 | 50.0 | 244 | 14 | Q52337 | CARSID PROTEIN (FRAME | 1.47e-01 |
| 35 | 53 | 50.0 | 338 | 11 | P70199 | ERF1 (FRAGMENT) | 1.47e-01 |
| 36 | 53 | 50.0 | 391 | 2 | P72859 | ASPARTATE TRANSAMINASE | 1.47e-01 |
| 37 | 53 | 50.0 | 544 | 2 | Q45145 | NEURAMINIDASE PRECURS | 1.47e-01 |
| 38 | 53 | 50.0 | 551 | 5 | Q96252 | ATP SYNTHASE ALPHA CHA | 1.47e-01 |
| 39 | 53 | 50.0 | 723 | 2 | Q92NM8 | PROLYL ENDOPEPTIDASE P | 1.47e-01 |
| 40 | 53 | 50.0 | 999 | 3 | Q90338 | SIMILARITY TO DNA POLY | 1.47e-01 |
| 41 | 53 | 50.0 | 2212 | 4 | Q93072 | MYELOBLAST KIAA0346 PR | 1.47e-01 |
| 42 | 52 | 49.1 | 287 | 10 | Q92V50 | MSF3.6 PROTEIN | 2.32e-01 |
| 43 | 52 | 49.1 | 355 | 8 | Q37447 | NADH DEHYDROGENASE SUB | 2.32e-01 |
| 44 | 52 | 49.1 | 457 | 5 | Q09579 | HYPOTHETICAL 52.3 KD P | 2.32e-01 |
| 45 | 52 | 49.1 | 924 | 5 | Q22239 | F40F9.6 PROTEIN | 2.32e-01 |

ALIGNMENTS

RESULT 1
ID Q92NJ7 PRELIMINARY: PRT: 367 AA.
AC Q92NJ7
DI 01-MAY-1999 (TRENBLREL. 10, Created)
DT 01-MAY-1999 (TRENBLREL. 10, Last sequence update)
DI 01-NOV-1999 (TRENBLREL. 12, Last annotation update)
DE POLLEN MAJOR ALLERGEN 1-2
OS Juniperus ashei (Ozark white cedar)
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC euphyllophytes; Spermatophyta; Coniferopsida; Coniferales;
CC Taxodiaceae; Juniperus
RN [1]
RP SEQUENCE FROM N.A.
RA MIDORO-HORIUTI T.M., GOLDBLUM R.M., KURSKY A., WOOD T.G.
RA BROOKS S.G.
RT Molecular cloning of polystatin cedar (Juniperus ashei) pollen allerg
RT allergen, JMAA 1.1
RE S. Juniperus (New York) to the USA/Juniperus ashei Juniperus
DR EMBL AF015543 AAC (1997)
DR EMBL AF015543 AAC (1997)
DR MENDEL: 36544; Juncas:1085136544
DR MENDEL: 36545; Juncas:1085136545
SQ SEQUENCE 367 AA: 39824 MW: 4228630 CRF32:

Query Match: 100.0% Score 106: DB 10: Length 467
Best Local Similarity 100.0% Pred. No.: 02e-12
Matches 15: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

DB 232 KSMKVIVAFNQFGPN 246
||| ||||| |||||
QY 1 KS KKVIVAFNQFGPN 15

RESULT 2
ID Q96395 PRELIMINARY: PRT: 375 AA.
AC Q96395
DI 01-FEB-1997 (TRENBLREL. 02, Created)
DT 01-FEB-1997 (TRENBLREL. 02, Last sequence update)
DI 01-NOV-1999 (TRENBLREL. 12, Last annotation update)
DE CHAO1 PRECURSOR
OS Chamaecyparis obtusa
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
CC euphyllophytes; Spermatophyta; Coniferopsida; Coniferales;
CC Taxodiaceae; Chamaecyparis.


```

GN PEL.
GS Musa acuminata (Banana).
CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida;
CC Zingiberales; Musaceae; Musa.
RN [1]
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, DWARF CAVENDISH; TISSUE=RIPE FRUIT;
RA DOMINGUEZ-PUIGJANER E.;
RI Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
RN [1]
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, DWARF CAVENDISH; TISSUE=RIPE FRUIT;
RA DOMINGUEZ-PUIGJANER E.;
RI Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: ELIMINATIVE CLEAVAGE OF PECTATE TO GIVE
CC OLIGOSACCHARIDES WITH 4-DEOXY-ALPHA-D-GLUC-4-ENURONOSYL GROUPS AT
CC THEIR NON-REDUCING ENDS.
DR EMBL: X92943; CAA63496.1;
DR MENDEL: 9136; Musac; 1089; 9136.
DR PFAM: PF00544; pec_lyase; 1.
KW Lyase.
FT SIGNAL: 1 20 P-TECTAL.
FT CHAIN: 21 401 PECTATE LYASE.
SQ SEQUENCE 396 AA: 43712 MW: 5598932A CRC32:
Query Match 77.4% Score 82; DB 10; Length 398;
Best Local Similarity 66.7% Pred. No. 2.87e-06;
Matches 10; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

DB 258 KMKQVTIAFNHFG 272
QY 1 KSMKVTVAFNQFG 15

RESULT 7 PRELIMINARY: PRT: 226 AA.
ID Q23657
AC Q23657;
DI 01-JAN-1998 (TRENBLrel. 05, Created)
DI 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DI 01-NOV-1999 (TRENBLrel. 12, Last annotation update)
DE PUTATIVE PECTATE LYASE (FRAGMENT).
GN A10.
CS Arabidopsis thaliana (Mouse-ear cress);
CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
CC core eudicots; Rosidae; eustosids II; Brassicales; Brassicaceae;
CC Arabidopsis.
RN [1]
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE: 97422403.
RA KULIKAUSSAS R., MCCORMICK S.;
RI "Identification of the tobacco and Arabidopsis homologues of the
RI pollen-expressed LAT59 gene of tomato.";
RI Plant Mol. Biol. 34:809-814(1997).
DR EMBL: U83622; AAB69761.1;
DR MENDEL: 25561; Arabid; 1088; 25561.
DR PFAM: PF00544; pec_lyase; 1.
KW Lyase.
FT NON_TER 1 1
FT NON_TER 226 226
SQ SEQUENCE 226 AA: 25103 MW: 5FLC8DD2 CRC32:
Query Match 73.6% Score 78; DB 10; Length 226;
Best Local Similarity 69.2% Pred. No. 2.96e-05;
Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

DB 196 KMKQVTIAFNHFG 208
QY 1 KSMKVTVAFNQFG 13

RESULT 8 PRELIMINARY: PRT: 401 AA.
ID Q24554
AC Q24554;
DI 01-JAN-1998 (TRENBLrel. 05, Created)
DI 01-JAN-1998 (TRENBLrel. 05, Last sequence update)
DI 01-NOV-1999 (TRENBLrel. 12, Last annotation update)
DE PECTATE LYASE PRECURSOR.
CC Zinnia elegans.
CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
CC core eudicots; Asterales; eustosids II; Asterales; Asteraceae;
CC Asteroideae; Heliantheae; Zinnia.
RN [1]
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE: 98345940.
RA DOMINGO C., ROBERTS K., STALEY N.J., CONNERTON J., MCCANN M.C.;
RI "A pectate lyase from Zinnia elegans is auxin inducible.";
RI Plant J. 13:17-28(1998);
DR EMBL: Y09341; CAA70735.1;
DR MENDEL: 26064; Zinel; 1088; 26064.
DR PFAM: PF00544; pec_lyase; 1.
DR PRINTS: PR00807; AMBALLERGEN.
KW Signal; Lyase.
FT SIGNAL: 1 20 P-TECTAL.
FT CHAIN: 21 401 PECTATE LYASE.
SQ SEQUENCE 401 AA: 44407 MW: C4C0F131 CRC32:
Query Match 73.6% Score 78; DB 10; Length 401;
Best Local Similarity 69.2% Pred. No. 2.96e-05;
Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

DB 261 KMKQVTIAFNHFG 273
QY 1 KSMKVTVAFNQFG 13

RESULT 9 PRELIMINARY: PRT: 455 AA.
ID Q64510
AC Q64510;
DI 01-AUG-1998 (TRENBLrel. 07, Created)
DI 01-AUG-1998 (TRENBLrel. 07, Last sequence update)
DI 01-NOV-1999 (TRENBLrel. 12, Last annotation update)
DE PUTATIVE PECTATE LYASE.
GN T20F6.14.
CS Arabidopsis thaliana (Mouse-ear cress);
CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
CC core eudicots; Rosidae; eustosids II; Brassicales; Brassicaceae;
CC Arabidopsis.
RN [1]
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA;
RA ROUNSLEY S.D., LIN X., KETCHUM K.A., CROSBY M.L., BRANDON R.C.,
RA SYKES S.M., KAUL S., MASON T.M., KERLAVAGE A.R., ADAMS M.D.;
RI SOMERVILLE C.R., VENTER J.C.;
RI Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC002521; AAC05350.1;
DR MENDEL: 28505; Arabid; 1088; 28505.
DR PFAM: PF00544; pec_lyase; 1.
DR PRINTS: PR00807; AMBALLERGEN.
KW Lyase.
FT NON_TER 1 1
FT NON_TER 455 455
SQ SEQUENCE 455 AA: 51257 MW: EBA0B82D CRC32:
Query Match 71.7% Score 76; DB 10; Length 455;
Best Local Similarity 69.2% Pred. No. 9.30e-05;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

DB 309 KMKQVTIAFNHFG 321
QY 1 KSMKVTVAFNQFG 13

RESULT 10 PRELIMINARY: PRT: 459 AA.
ID Q23665
AC Q23665;

```

```

SQ      SEQUENCE   431 AA:  47770 MW: 868EC285 CAC32:

Query Match          70.8%: Score 75; DB 10; Length 431;
Best Local Similarity 61.5%: Pref. NO. 1.64e-04;
Matches            8; Conservative 4; Mismatches 1; Indels 0; Gaps 5

Db       289 KAMQVTIAYNHF5 301
        ||| ||| |||| 
Qy       1 KMKKVIVAFNQF5 13

RESULT    12
ID         QC4084
AC         QC4084:
DT         01-NOV-1996 ((REMBLrel_01; created))
DI         01-NOV-1996 ((REMBLrel_01; last sequence update))
DT         01-NOV-1996 ((REMBLrel_12; last annotation update))
DE         PEPTATE LYASE (EC 4.2.2.2) (FRAGMENT).
OS         Nicotiana tabacum (Common tobacco).
OC         Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophytes;
OC         euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC         core eudicots; Asteridae; euasterids II; Solanales; Solanaceae.
OC         Nicotiana.
RN         111
```

RP SEQUENCE FROM N.A.
RC STRA-N-SAMSON;
RA ROGERS H.C., LONSDALE S.M.;

| | | |
|----|----------|---|
| CC | 1-1 | CATALYTIC ACTIVITY: ELIMINATIVE CLEAVAGE OF PEPTATE TO GIVE |
| CC | 1-1 | OLIGOSACCHARIDES WITH 4-DEOXY-ALPHA-D-GULO-4-ENOKRONYL GROUPS AND |
| CC | 1-1 | THEIR NON-REDUCING ENDS |
| CC | 1-1 | DEVELOPMENTAL STAGE: EXPRESSED DURING MICHROSPOROGENESIS |
| CC | 1-1 | SIMILARITY: TO TOMATO PROTEIN P56 PRECURSOR (AC P15721) |
| DR | EMBL: | X61100; CA443413.1 |
| DR | MEDLINE: | 16497; NICTD:1088;1649 |
| DR | PIRANK: | PF00544; pcc-lyase; i. |
| DR | LYASE: | |
| KW | NON-TER | i |
| FT | SEQUENCE | 263 AA; 29449 MW. -24305A CRG32. |
| SQ | | |

```
Query Match      99.06%    Seq ID     Length   2873
Best Local Similarity  51.58%    P-Val NO. 2.8E+04
Matches          40 Mismatches 1 Indels  50 Gaps

DB       136 KMKITZAYNES 148
              111111111
CV        1 KSKKIVARNS 14
```

```
RESULT         13
ID Q40319 PRELIMINARY: FRT: 450 AA.
AC Q40319:
DT 01-NOV-1996 (TRENDELIN) 01, Created
DT 01-NOV-1996 (TRENDELIN) 01, Last sequence update)
```

DI NOV-1993 [JHEMDEL: 127, EST ANNOTATION UPDATE]
 DE PECTATE LYASE, HOMOLOG.
 CS Medicago sativa (Altaifta).
 CS Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophytes:
 OC euphyllophytes: Spermatophyta: Magnoliophyta: eudicotyledons:
 OC core eudicots: Rosidae: eurosids I: Fabales: Fabaceae: Papilionaceae:
 OC Medicago.
 RC [1]
 RN SEQUENCE FROM N.A.
 RA W3.Y. QIU X. DE S., ERICKSON L.
 RP Submitted (NOV-1995) to the EMBL/GenBank/DBSJ databases.
 RL EMBL: G41472; AAA86441.1; ...
 DR MENDEL: 9042; Medsat:098;9042.
 CR PRAM: PFC0544; pec-lyase.1.
 CR PRINTS: PR00867; AMBALLERGEN.
 CR Lyase.
 SQ SEQUENCE 450 AA: 50155 MW: B35FFCF9E CRC32:
 Query Match: 69.8% Score 74. DR 10: Length 450:

Query Match: 69.88; Score 74; DB 10; Length 450;

A B E F H
I J K L M N O P Q R S T U V W X Y Z
- - - - -
(TM)

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MFstich_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Mon Jun 19 16:12:29 2000: MasPar time 3.59 Seconds
Tabular output not generated.
210,886 Million cell updates/sec

Title: >US-09-142-524A-10
Description: (1,32) from US09142524A.pep
Perfect Score: 427
Sequence: 1 EYLLSARDVLAVYSKRMKVIVAFNFGPN 32

Scoring table: PAM 150
Gap 11

Searched: 189963 seqs, 23686106 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-geneseq35
l-geneseqp

Statistics: Mean 23.654: Variance 92.032 scale 0.257

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

| SUMMARY | | | | SUMMARY | | | |
|------------|-------|-----------------------|-----|-------------------------|-----------|------------|-------|
| Result No. | Score | Query Match Length DB | ID | Description | Pred. No. | Result No. | Score |
| 1 | 114 | 50.2 | 93 | M. leprae GroES-like p | 8.77e-03 | 1 | 114 |
| 2 | 103 | 45.4 | 102 | GroES structural prote | 5.04e-02 | 2 | 103 |
| 3 | 98 | 43.2 | 15 | Japanese cedar pollen | 2.57e-01 | 3 | 98 |
| 4 | 98 | 43.2 | 15 | Residues 211-225 of Cr | 2.57e-01 | 4 | 98 |
| 5 | 98 | 43.2 | 17 | Sugi allergen protein | 2.57e-01 | 5 | 98 |
| 6 | 98 | 43.2 | 20 | T-cell epitope peptide | 2.57e-01 | 6 | 98 |
| 7 | 98 | 43.2 | 20 | Cry j I Japanese Cedar | 2.57e-01 | 7 | 98 |
| 8 | 98 | 43.2 | 20 | Cry j I pollen allergen | 2.57e-01 | 8 | 98 |
| 9 | 98 | 43.2 | 26 | Cry j I pollen allergen | 2.57e-01 | 9 | 98 |
| 10 | 98 | 43.2 | 26 | Cry j I pollen allergen | 2.57e-01 | 10 | 98 |
| 11 | 98 | 43.2 | 26 | Cry j I pollen allergen | 2.57e-01 | 11 | 98 |
| 12 | 98 | 43.2 | 28 | Cry j I pollen allergen | 2.57e-01 | 12 | 98 |
| 13 | 98 | 43.2 | 28 | Cry j I pollen allergen | 2.57e-01 | 13 | 98 |
| 14 | 98 | 43.2 | 28 | Cry j I pollen allergen | 2.57e-01 | 14 | 98 |
| 15 | 98 | 43.2 | 30 | Cry j I pollen allergen | 2.57e-01 | 15 | 98 |
| 16 | 98 | 43.2 | 30 | T-cell epitope peptide | 2.57e-01 | 16 | 98 |
| 17 | 98 | 43.2 | 30 | Cry j I pollen allergen | 2.57e-01 | 17 | 98 |
| 18 | 98 | 43.2 | 36 | Cry j I pollen allergen | 2.57e-01 | 18 | 98 |
| 19 | 98 | 43.2 | 353 | Japanese cedar pollen | 2.57e-01 | 19 | 98 |
| 20 | 98 | 43.2 | 353 | Cedar pollen allergen | 2.57e-01 | 20 | 98 |
| 21 | 98 | 43.2 | 354 | Japanese cypress polle | 2.57e-01 | 21 | 98 |
| 22 | 98 | 43.2 | 354 | Chamaecyparis obtusa p | 2.57e-01 | 22 | 98 |
| 23 | 98 | 43.2 | 367 | Jun S I. | 2.57e-01 | 23 | 98 |

| | | | | | | | |
|----|----|------|------|---|--------|-------------------------|----------|
| 24 | 98 | 43.2 | 374 | 1 | R45541 | Cry j I pollen allergen | 2.57e-01 |
| 25 | 98 | 43.2 | 374 | 1 | R31937 | Cry j I. | 2.57e-01 |
| 26 | 98 | 43.2 | 374 | 1 | R82450 | Cry j I Japanese Cedar | 2.57e-01 |
| 27 | 98 | 43.2 | 374 | 1 | R60166 | Japanese cedar pollen | 2.57e-01 |
| 28 | 98 | 43.2 | 375 | 1 | W04345 | Chamaecyparis obtusa p | 2.57e-01 |
| 29 | 95 | 41.9 | 80 | 1 | W27369 | Multi-epitope peptide | 4.77e-01 |
| 30 | 95 | 41.9 | 135 | 1 | W27180 | Multi-epitope peptide | 4.77e-01 |
| 31 | 95 | 41.9 | 134 | 1 | W27171 | Multi-epitope peptide | 4.77e-01 |
| 32 | 94 | 41.4 | 370 | 1 | R4378 | Jun V I. | 5.84e-01 |
| 33 | 92 | 40.5 | 81 | 1 | W8358 | Sugi allergen protein | 8.83e-01 |
| 34 | 90 | 39.5 | 13 | 1 | Y6196 | Peptide used for the t | 1.33e-00 |
| 35 | 88 | 37.9 | 47 | 1 | W80357 | Sugi allergen protein | 2.98e-00 |
| 36 | 88 | 37.9 | 47 | 1 | R80351 | Sugi allergen protein | 2.98e-00 |
| 37 | 85 | 37.4 | 22 | 1 | R92179 | Cry j I Japanese Cedar | 4.45e-00 |
| 38 | 84 | 37.0 | 47 | 1 | W80352 | Sugi allergen protein | 4.45e-00 |
| 39 | 82 | 36.1 | 11 | 1 | Y01795 | Peptide used for the t | 6.63e-00 |
| 40 | 82 | 36.1 | 47 | 1 | W80356 | Sugi allergen protein | 6.63e-00 |
| 41 | 82 | 36.1 | 47 | 1 | W80357 | Sugi allergen protein | 6.63e-00 |
| 42 | 82 | 36.1 | 47 | 1 | W80357 | Sugi allergen prtein | 6.63e-00 |
| 43 | 79 | 34.8 | 99 | 1 | RC5700 | MPB-57 protein | 1.22e-01 |
| 44 | 79 | 34.8 | 869 | 1 | R53732 | S. cerevisiae P.c1 pro | 1.22e-01 |
| 45 | 74 | 32.6 | 1021 | 1 | W70526 | Human thymus receptor | 3.17e-01 |

AL COMMENTS

RESULT 1
ID R67386 standard; Protein: 93 AA.
AC R67386;
DI 22-JUN-1995 (first entry)
DE M. leprae GroES-like protein
KW Urease; immunogen; vaccine; diagnostic; heat shock protein; HSP;
KW GroES-like protein; Helicobacter felis.
OS Mycobacterium leprae.
PN W09426901-A.
PD 24-NOV-1994.
PF 19-MAY-1994; E01625.
PR 19-MAY-1993; EP-401309.
PR 19-NOV-1993; WO-E03259.
PA (INRM) INST NAT SANTE & RECH MEDICALE
PA (INSP) INST PASTEUR.
PI Ferrero R., Labigne A., Suerbaum S., Thiberge J.
DR WPI: 95-008797/01.
PT DNA from Helicobacter pylori and Helicobacter felis - used to
PT develop prods. for detection, treatment and prevention of
PT Helicobacter infection
PS Disclosure: Fig. 7B(i-ii); 158pp; English.
CC The sequence of the Helicobacter pylori heat shock protein B
CC (given as R67373) was compared to that of other GroES-like
CC proteins from Legionella pneumophila (R67387), Escherichia coli
CC (R67390), Clostridium perfringens (R67397), Mycobacterium leprae
CC (R67386) and thermophilic bacterium (R67388), and regions
CC of homology were identified.
SQ Sequence 93 AA;

Query Match 50.2%; Score 114; DB 1; Length 93;
Best Local Similarity 100.0%; Pred. No. 8.77e-03;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 77 EYLLSARDVLAVYSK 93
QY 1 EYLLSARDVLAVYSK 17

RESULT 2
ID R22364 standard; Protein: 102 AA.
AC R22364;
DI 29-JUL-1992 (first entry)
DE GroES structural protein.
KW Heat shock protein; groES gene.
OS Streptomyces albus.
PN W09204452-A.
PD 19-MAR-1992.

PS 03-SEP-1991: F00701.
 PR 10-SEP-1990: FR-01186.
 PA (INSP) INST. PASTEUR.
 PI Mazodier P, Guellieff M G.
 PR WPI: 92-114/68/14.
 PR NSDB: 022483.
 PR Recombinant DNA cont. heat inducible promoter and heterologous
 PI gene - also vectors, transformed cells and new heat shock
 PI proteins of Streptococcus albus
 PR Disclosure: P1a 5: 50pp: French.
 CC The sequence is that of the GROS protein which is encoded by the
 CC structural gene gros. See also Q22477-022486.
 SQ Sequence 102 AA:

Query Match 45.4% Score 103: DB 1: Length 102:
 Best Local Similarity 76.5%: Pred. No. 9,04e-02:
 Matches 13: Conservative 3: Mismatches 1: Indels 0: Gaps 0:

DB 86 EYVLVSARDVLAIIEK 102
 QY 1 EYVLVSARDVLAVVSK 17
 : ||||| |||||

RESULT 3
 ID R89293 standard: peptide: 15 AA.

AC R89293:
 DT 12-MAR-1996 (first entry)
 DE Japanese cedar pollen allergen Cry I derived T-cell epitope peptide.
 KW Japanese cedar: pollen allergen: Cry I: T-cell epitope: peptides:
 KW prevention: treatment: cryptomeria pollenosis.
 CS Cryptomeria japonica.
 PN J07118295-A.
 PD 09-MAY-1995.
 PR 20-OCT-1993: 262626.
 PR 20-OCT-1993: JP-262626.
 PA (MEIP) MEIJ MILK PROD CO LTD.
 DR WPI: 95-203834/27.
 PI New cryptomeria pollen allergen T-cell epitope peptide - used for
 PI prevention, treatment and investigation of Japanese cedar pollenosis
 PS Claim 5: Page 2: 8pp: Japanese.
 CC R75388 is the Japanese cedar pollen allergen Cry I, from which the
 CC T-cell epitope peptides R89289-R89295 were derived. The peptides
 CC can be used for the prevention and treatment of cryptomeria
 CC pollenosis, and also for the investigation of pollenosis.
 SQ Sequence 15 AA:

Query Match 43.2% Score 98: DB 1: Length 15:
 Best Local Similarity 86.7%: Pred. No. 2,57e-01:
 Matches 13: Conservative 1: Mismatches 1: Indels 0: Gaps 0:

DB 1 KSMKVTVAFNQFGPN 15
 QY 18 RRMKVTVAFNQFGPN 32
 : ||||| |||||

RESULT 4

ID W57755 standard: peptide: 15 AA.

AC W57755:
 DT 17-SEP-1996 (first entry)
 DE Residues 211-225 of Cry j 1.
 KW Cry j 1: Japanese cedar pollen antigen: allergy: immunotherapy:
 KW HLA class II molecule.
 CS Cryptomeria japonica.
 PN W09820902-A1.
 PD 22-MAY-1996.
 PR 12-NOV-1997: J04129.
 PR 13-NOV-1996: JP-302053.
 PA (MEIP) MEIJ MILK PROD CO LTD.
 PI Cairiki K, Kuno K, Kume A, Sone T.
 DR WPI: 98-29767/26.
 PI Peptides derived from Japanese cedar pollen antigens are
 PI immunotherapeutic agents - useful for allergy treatment and typing
 PI HLA class II molecules in allergy sufferers

PS Claim 12: Page 28: 50pp: Japanese.
 CC This sequence represents residues 211-225 of the Cry j 1 protein, and
 CC is a peptide of the invention. The peptides are derived from Japanese
 CC cedar pollen antigens, and are used as immunotherapeutic agents in the
 CC treatment of allergy. The peptides can be used for identification and
 CC typing of the particular HLA class II molecules in an allergy sufferer,
 CC and also for peptide immunotherapy of an allergy. Using these peptides
 CC the immunotherapy can be targeted more specifically to the requirements
 CC of the individual patient, allowing more effective treatment of an
 CC allergy, including those patients for whom treatment with a conventional
 CC immunotherapeutic agent is ineffective.
 SQ Sequence 15 AA:

Query Match 43.2% Score 98: DB 1: Length 15:
 Best Local Similarity 86.7%: Pred. No. 2,57e-01:
 Matches 13: Conservative 1: Mismatches 1: Indels 0: Gaps 0:

DB 1 KSMKVTVAFNQFGPN 15
 QY 18 RRMKVTVAFNQFGPN 32
 : ||||| |||||

RESULT 5

ID W80349 standard: peptide: 17 AA.

AC W80349:
 DT 11-JAN-1999 (first entry)
 DE Sugi allergen protein Cryj2 derived epitope for T cells.
 KW T cell epitope: sugi allergen proteins Cryj1: Cryj2: treatment:
 KW sugi-pollenosis: allergic reaction: pollen.
 CS Synthetic.
 PN J0259198-A.
 PD 29-SEP-1998.
 PR 22-DEC-1997: 353448.
 PR 24-DEC-1996: JP-34341.
 PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
 PA (SANY) SANKYO CO LTD.
 DR WPI: 98-577037/49.
 PI A linked T cell epitope peptide - used for the treatment of
 PI sugi-pollenosis
 PS Claim 7: Page 18: 21pp: Japanese.
 CC W80339-58 represent epitopes for T cells, derived from the sugi allergen
 CC proteins Cryj1 (W80339-44, W80350-53 and W80356-58) and Cryj2 (W80345-49
 CC and W80354-55). The peptides are useful for the treatment of
 CC sugi-pollenosis, an allergic reaction of the body to pollen.
 SQ Sequence 17 AA:

Query Match 43.2% Score 98: DB 1: Length 17:
 Best Local Similarity 85.7%: Pred. No. 2,57e-01:
 Matches 13: Conservative 1: Mismatches 1: Indels 0: Gaps 0:

DB 2 KSMKVTVAFNQFGPN 16
 QY 18 RRMKVTVAFNQFGPN 32
 : ||||| |||||

RESULT 6

ID W42144 standard: peptide: 20 AA.

AC W42144:
 DT 16-JUN-1998 (first entry)
 DE T-cell epitope peptide 24 from Japanese cypress pollen antigen: Chao1.
 KW Japanese cypress pollen: antigen: T-cell epitope: Chao1: Chao2.
 KW diagnosis: allergy: spring tree pollen disease: pollenosis:
 CS Chamaecyparis obtusa.
 PN W09747548-A1.
 PD 18-DEC-1997.
 PR 12-JUN-1997: J02031.
 PR 14-JUN-1996: JP-153527.
 PA (MEIP) MEIJ MILK PROD CO LTD.
 PI Dairiki K, Kuno K.
 DR WPI: 98-052242/05.
 PI T-cell epitope peptide portion of Japanese cypress pollen antigens
 PI Chao1 and Chao2 - used for diagnosis and treatment of spring tree
 PI pollen disease

PS Claim 1: Page 27-28: 71pp: Japanese.
CC The present sequence represents a T-cell epitope peptide from Japanese
CC cypress pollen antigen Chaoi. The present invention describes peptides
CC which correspond to the T-cell epitope sites on Japanese cypress pollen
CC antigens Chaoi and Chao2. The peptides can be used as a reagent for the
CC diagnosis of allergy to Japanese cypress pollen, and as an antigen in
CC the treatment and prevention of spring tree pollen disease in which the
CC pollinosis involves reactivity to Japanese cypress pollen.
SQ Sequence 20 AA:

Query Match 43.2% Score 98: DB 1: Length 20:
Best Local Similarity 86.7%: Pred. No. 2,578-01:
Matches 13: Conservative 1: Mismatches 1: Indels 0: Gaps 0:

DB 1 KSMKVTVAFNFGPN 15
QY 18 RRMKVTVAFNFGPN 32

RESULT 7
ID R82512 standard: Protein: 20 AA.
AC R82512:
DE 15-APR-1996 (first entry)
DE Cry j 1 Japanese cedar pollen allergen peptide fragment (CJI-22):
KW Cry j 1: Japanese cedar pollen allergen; modified: drug production:
KW allergy: Cryptomeria japonica.
OS Cryptomeria japonica.
PN WO9527786-A1
PD 19-OCT-1995
PF 06-APR-1995: U04245.
PR 08-APR-1994: US-226248.
PR 06-DEC-1994: US-350225.
PA (IMMU-) IMMUNOLOGIC PHARM CORP.
PI Chen X, Evans S, Franzen HM, Kuo M, Powers SP:
PI Skated 2:
PI WPI: 95-366391/47.
DR Modified Cryptomeria japonica (Cry j) 1 peptide(s) - useful for
PI treating allergy to Japanese cedar pollen allergen or
PI immunologically cross reactive allergens
PS Disclosure: Figure 2: 60pp: English.
CC Novel peptides of cry j 1 have been modified as a part of a
CC preformulation scheme to develop an optimised drug product for
CC therapeutic treatment of humans suffering from allergy to Japanese
CC cedar pollen allergen of an allergen which is immunologically cross
CC reactive with Japanese cedar pollen allergen. Such modified peptides
CC possess certain characteristics which render them particularly
CC suitable for drug product formulation. Peptide fragments of Cry j 1,
CC modified and unmodified, are given in P82491-R82525. This peptide
CC fragment corresponds to amino acids 211-230 of the allergen mature
CC protein.
SQ Sequence 20 AA:

Query Match 43.2% Score 98: DB 1: Length 20:
Best Local Similarity 86.7%: Pred. No. 2,578-01:
Matches 13: Conservative 1: Mismatches 1: Indels 0: Gaps 0:

DB 1 KSMKVTVAFNFGPN 15
QY 18 RRMKVTVAFNFGPN 32

RESULT 8
ID R45563 standard: Protein: 20 AA.
AC R45563:
DE 13-JUL-1994 (first entry)
DE Cry j 1 pollen allergen peptide CJI-22.
KW Japanese cedar: detection: allergy; treatment: diagnosis;
KW T cell epitope: sensitivity.
OS Cryptomeria japonica.
PN WO9401540-A.
PD 20-JAN-1994.
PF 15-JAN-1993: U00139.
PR 10-JUL-1992: WO-905661.

PR 01-SEP-1992: US-938990.
PA (IMMU-) IMMUNOLOGIC PHARM CORP.
PI Bond JF, Garman RD, Griffith LJ, Kuo M, Pollock J:
PI WPI: 94-035066/04.
DE Antigen derived from Japanese cedar pollen allergen Cry j 1
DE contain at least two T cell epitope(s), used to treat or diagnose
DE allergy
PS Claim 1: Fig 13: 137pp: English.
CC The sequence is that of an isolated peptide of the Japanese cedar
CC pollen allergen Cry j 1 (amino acids 211-230). The peptide, CJI-22,
CC can be used for the treatment and diagnosis of allergies associated
CC with Japanese cedar pollen. It has enhanced therapeutic properties
CC but reduced side effects compared to naturally occurring allergens
SQ Sequence 20 AA:

Query Match 43.2% Score 98: DB 1: Length 20:
Best Local Similarity 86.7%: Pred. No. 2,578-01:
Matches 13: Conservative 1: Mismatches 1: Indels 0: Gaps 0:

DB 1 KSMKVTVAFNFGPN 15
QY 18 RRMKVTVAFNFGPN 32

RESULT 9
ID R45592 standard: Protein: 26 AA.
AC R45592:
DE 13-JUL-1994 (first entry)
DE Cry j 1 pollen allergen peptide CJI-43.10.
KW Japanese cedar: detection: allergy; treatment: diagnosis;
KW T cell epitope: sensitivity.
OS Cryptomeria japonica.
PN WO9401560-A.
PD 20-JAN-1994.
PF 15-JAN-1993: U00139.
PR 10-JUL-1992: WO-905661.
PR 01-SEP-1992: US-938990.
PA (IMMU-) IMMUNOLOGIC PHARM CORP.
PI Bond JF, Garman RD, Griffith LJ, Kuo M, Pollock J:
PI WPI: 94-03 066/04.
DE Antigen derived from Japanese cedar pollen allergen Cry j 1
DE contain at least two T cell epitope(s), used to treat or diagnose
DE allergy
PS Claim 76: Fig 18: 137pp: English.
CC The sequence is that of an isolated peptide of the Japanese cedar
CC pollen allergen Cry j 1. The peptide, CJI-43.10, can be used for
CC the treatment and diagnosis of allergies associated with Japanese
CC cedar pollen. It has enhanced therapeutic properties but reduced
CC side effects compared to naturally occurring allergens.
SQ Sequence 26 AA:

Query Match 43.2% Score 98: DB 1: Length 26:
Best Local Similarity 86.7%: Pred. No. 2,578-01:
Matches 13: Conservative 1: Mismatches 1: Indels 0: Gaps 0:

DB 1 KSMKVTVAFNFGPN 15
QY 18 RRMKVTVAFNFGPN 32

RESULT 10
ID R45594 standard: Protein: 26 AA.
AC R45594:
DE 13-JUL-1994 (first entry)
DE Cry j 1 pollen allergen peptide CJI-43.12.
KW Japanese cedar: detection: allergy; treatment: diagnosis;
KW T cell epitope: sensitivity.
OS Cryptomeria japonica.
PN WO9401560-A.
PD 20-JAN-1994.
PF 15-JAN-1993: U00139.
PR 10-JUL-1992: WO-905661.
PR 01-SEP-1992: US-938990.

PI Bond JF, Garman RD, Griffith IJ, Kuo M, Pollock J:
 DPI WPI: 94-035066/04.
 DR Antigen(s) derived from Japanese cedar pollen allergen Cry j 1.
 PI Antigen(s) derived from Japanese cedar pollen allergen Cry j 1.
 PI contain at least two I cell epitope(s), used to treat or diagnose
 PI allergy.
 PS Claim 76: Fig 18: 137pp: English.
 CC The sequence is that of an isolated peptide of the Japanese cedar
 CC pollen allergen Cry j 1. The peptide, CJI-43.8, can be used for
 CC the treatment and diagnosis of allergies associated with Japanese
 CC cedar pollen. It has enhanced therapeutic properties but reduced
 CC side effects compared to naturally occurring allergens.
 SQ Sequence 28 AA:

Query Match 43.28; Score 98; DB 1; Length 28;
Best Local Similarity 86.7%; Pred No 2,576-91;
Matches 13; Conservative 1; Mismatches 1; Indels

```

1 KSMKVTVAFNQFCPN 15
: : : : : : : : : :
18 RRMKVTVAFNQFCPN 32

```

RESULT 13

R45589 standard: Protein: 28 AA.

AC R45589;

AD 13-JUL-1994 (first entry)

CE Cry j 1 pollen allergen peptide CRI-43.7.

DE Japanese cedar: detection: allergy; treatment: diagnosis.

KW T cell epitope; sensitivity.

KW Cry: *Prunella japonica*.

OS WC9401360-A.

PN 20-JAN-1994.

PP 15-JAN-1993: U00139.

PE 10-JUL-1992: WC-U05661.

PR 01-SEP-1992: CS-938990.

PA (IMMU-) IMMUNOLOGIC PHARM. CORP.

PI Bond JF, Garman RD, Griffin JJ, Kio M, Pollock JI.

PT WPI: 94-035065/04.

PT Antigen(s) derived from Japanese cedar pollen allergen Cry j 1.

PT contain at least two T cell epitope(s), used to treat or diagnose allergy.

PS Claim 36; Fig 18; 337pp; Eng. 1.

PS The sequence is that of an isolated peptide of the Japanese cedar

CC pollen allergen Cry j 1. The peptide, CRI-43.7, can be used for

CC the treatment and diagnosis of allergies associated with Japanese

CC cedar pollen. It has enhanced therapeutic properties but reduced

CC side effects compared to naturally occurring allergens.

CC Sequence 28 AA.

SQ

| | | | | |
|-----------------------|-------|---------------|--------|-----------|
| Query Match | 43.2% | Score 34 | (p. 1) | Count 28 |
| Best Local Similarity | 96.7% | Prod. No. 276 | | |
| Matches | 13 | Conservative | 1 | Matches 3 |
| | | | | Index 2 |

```

00 1 KSMKV:VAFNQ:GPN 15
    : ||| ||| |||
01 18 8RMKV:VAFNQ:GPN 32

```

| | |
|--------|--|
| RESULT | 14 |
| ID | R45593 standard; Protein: 28 AA. |
| AC | R45593: |
| DT | 13-JUL-1994 (first entry) |
| DE | Cry j I pollen allergen peptide 201-4311. |
| DE | Japanese cedar; detection: a.e.t.; treatment; diagnosis. |
| KW | T cell epitope; sensitivity. |
| KW | Cryptomeria japonica. |
| CS | CS |
| WC | WC9401580-A. |
| PN | PN |
| PD | 20-JAN-1994. |
| PF | 15-JAN-1993: U00139. |
| PR | 1C-JUL-1992: WS-U05661. |
| PR | 01-SEP-1992: WS-93899C. |
| PA | (IMMU-) IMMUNOLOG PHARM CORP. |
| PI | Bond JF, Garman RD, Griffith J, Kuo M, Pollock J. |

DR WPI: 94-035086/34.
 PI Antigens derived from Japanese cedar pollen allergen Cry j 1 -
 PI contain at least two T cell epitope(s), used to treat or diagnose
 PI allergy.
 PS Claim 76: Fig 18: 137pp: English.
 CC The sequence is that of an isolated peptide of the Japanese cedar
 CC pollen allergen Cry j 1. The peptide, CJI-43.11, can be used for
 CC the treatment and diagnosis of allergies associated with Japanese
 CC cedar pollen. It has enhanced therapeutic properties but reduced
 CC side effects compared to naturally occurring allergens.
 SQ Sequence 28 AA:

Query Match 43.2%, Score 98: DB 1: Length 26:
 Best Local Similarity 86.7%: Pred No: 2,578-01:
 Matches 13: Conservative 1: Mismatches 1: Indels 0: Gaps 0:

EE 1 KSMKVTVAFNQFGPN 15
 : |||: ||||: |||
 QY 18 PRMKVTVAFNQFGPN 32

RESULT 15
 ID R45588 standard: Protein: 10 AA.
 AC R45588:
 DE 13-JUL-1994 (first entry)
 DE Cry j 1 pollen allergen peptide CJI-43.6.
 KW Japanese cedar: detection: allergy: treatment: diagnosis:
 KW T cell epitope: sensitivity.
 QS Cryptomeria japonica
 PN NC0401560-A.
 PD 20-JAN-1994:
 PF 15-JAN-1993: U00139.
 PR 10-JUL-1992: W0-005661.
 PR 01-SEP-1992: US-93890.
 PA (IMMU-) IMMUNOLOGIC PHARM CORP.
 PI Bond JF, Garman RD, Griffith TJ, Kuo M, Pollock J:
 DR WPI: 94-035086/34.
 PI Antigens derived from Japanese cedar pollen allergen Cry j 1 -
 PI contain at least two T cell epitope(s), used to treat or diagnose
 PI allergy.
 PS Claim 76: Fig 18: 137pp: English.
 CC The sequence is that of an isolated peptide of the Japanese cedar
 CC pollen allergen Cry j 1. The peptide, CJI-43.6, can be used for
 CC the treatment and diagnosis of allergies associated with Japanese
 CC cedar pollen. It has enhanced therapeutic properties but reduced
 CC side effects compared to naturally occurring allergens.
 SQ Sequence 30 AA:

Query Match 43.2%, Score 98: DB 1: Length 30:
 Best Local Similarity 86.7%: Pred No: 2,578-01:
 Matches 13: Conservative 1: Mismatches 1: Indels 0: Gaps 0:

EE 1 KSMKVTVAFNQFGPN 15
 : |||: ||||: |||
 QY 18 PRMKVTVAFNQFGPN 32

Search completed: Mon Jun 19 16:12:39 2000
 Job time : 11 secs.

Release 3.1A John F. Collins, Biocomputing Research Unit,
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Tue Jun 20 13:36:17 2000: MasPar time 2.72 Seconds
tabular output not generated. 169,949 Million cell updates/sec

Title: >US-09-142-524A-10
Description: (1-32) from US09142524A.pep
Perfect Score: 227
Sequence: 1: EHYILSARDVLAVYSKRMKMTVAFNQFGPN 32
Scoring table: PAM 150
Gap 11
Searched: 145341 seqs. 14437480 residues
Post-processing: Minimum Match 0%
Listing first 45 summaries
Database: a-issued
1:5A_COMB 2:5B_COMB 3:6_COMB 4:PC1_COMB 5:backfiles:

Statistics: Mean 22.169; Variance 87.947; scale 0.252
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score the result being printed,
and is derived by analysis of the total scores distribution.
SUMMARIES

| Result No. | Score | Query Match | ID | Description | Prior No. |
|------------|-------|-------------|------|--------------------------|-----------|
| 1 | 114 | 50.2 | 93 | Sequence 36, Application | 4,560-03 |
| 2 | 72 | 31.7 | 2050 | Sequence 2, Application | 2,430-01 |
| 3 | 71 | 31.3 | 558 | Sequence 2, Application | 2,950-01 |
| 4 | 70 | 30.8 | 94 | Sequence 38, Application | 3,570-01 |
| 5 | 70 | 30.8 | 1895 | Sequence 4, Application | 3,570-01 |
| 6 | 70 | 30.8 | 2813 | Sequence 2, Application | 3,570-01 |
| 7 | 68 | 30.0 | 254 | Sequence 23, Application | 5,210-01 |
| 8 | 68 | 30.0 | 254 | Sequence 23, Application | 5,210-01 |
| 9 | 68 | 30.0 | 269 | Sequence 29, Application | 5,210-01 |
| 10 | 68 | 30.0 | 304 | Sequence 27, Application | 5,210-01 |
| 11 | 68 | 30.0 | 316 | Sequence 15, Application | 5,210-01 |
| 12 | 68 | 30.0 | 316 | Sequence 16, Application | 5,210-01 |
| 13 | 67 | 29.5 | 368 | Sequence 3, Application | 6,300-01 |
| 14 | 67 | 29.5 | 737 | Sequence 5, Application | 6,300-01 |
| 15 | 67 | 29.5 | 737 | Sequence 4, Application | 6,300-01 |
| 16 | 67 | 29.5 | 1687 | Sequence 29, Application | 6,300-01 |
| 17 | 67 | 29.5 | 1704 | Sequence 10, Application | 6,300-01 |
| 18 | 67 | 29.5 | 3387 | Sequence 2, Application | 7,600-01 |
| 19 | 66 | 29.1 | 1876 | Sequence 2, Application | 7,600-01 |
| 20 | 65 | 28.6 | 94 | Sequence 39, Application | 9,160-01 |
| 21 | 65 | 28.6 | 97 | Sequence 40, Application | 9,160-01 |
| 22 | 65 | 28.6 | 988 | Sequence 15, Application | 9,160-01 |
| 23 | 65 | 28.6 | 988 | Sequence 15, Application | 9,160-01 |

| RESULT ID | ID | US-08-467-822-36 | STANDARD | PRI | 93 AA |
|-----------|----|--|----------|-----|-------|
| XX | XX | XXXXXX | | | |
| XX | XX | | | | |
| XX | XX | | | | |
| DE | DE | Sequence 36, Application US/08467822 | | | |
| XX | XX | Sequence 36, Application US/08467822 | | | |
| CC | CC | Patent No. 5843450 | | | |
| CC | CC | GENERAL INFORMATION: | | | |
| CC | CC | APPLICANT: Lablone, Agnes | | | |
| CC | CC | APPLICANT: Sauerbaum, Sebastian | | | |
| CC | CC | APPLICANT: Ferrero, Richard L. | | | |
| CC | CC | APPLICANT: Tiberge, Jean-Michel | | | |
| CC | CC | TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST | | | |
| CC | CC | TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE | | | |
| CC | CC | TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENTS INCLUDING SAID | | | |
| CC | CC | TITLE OF INVENTION: POLYPEPTIDES | | | |
| CC | CC | NUMBER OF SEQUENTS: 46 | | | |
| CC | CC | CORRESPONDENT'S ADDRESS: | | | |
| CC | CC | ADDRESSER: Lablone, Agnes | | | |
| CC | CC | STREET: 1400 I Street, N.W. | | | |
| CC | CC | CITY: Washington | | | |
| CC | CC | STATE: D.C. | | | |
| CC | CC | COUNTRY: USA | | | |
| CC | CC | ZIP: 20005-3315 | | | |
| CC | CC | COMPUTER READABLE FORM: | | | |
| CC | CC | MEDIUM TYPE: Floppy disk | | | |
| CC | CC | COMPUTER: IBM PC compatible | | | |
| CC | CC | OPERATING SYSTEM: PC-DOS/MS-DOS | | | |
| CC | CC | SOFTWARE: Patent In Release #1.0, Version #1.30 | | | |
| CC | CC | CURRENT APPLICATION DATA: | | | |
| CC | CC | APPLICATION NUMBER: US/08/467.822 | | | |
| CC | CC | FILING DATE: 05-JUN-1995 | | | |
| CC | CC | CLASSIFICATION: 435 | | | |
| CC | CC | PRIOR APPLICATION DATA: | | | |
| CC | CC | APPLICATION NUMBER: US 08/447,177 | | | |
| CC | CC | FILING DATE: 19-MAY-1995 | | | |
| CC | CC | CLASSIFICATION: 435 | | | |
| CC | CC | PRIOR APPLICATION DATA: | | | |
| CC | CC | APPLICATION NUMBER: US 08/432,697 | | | |
| CC | CC | FILING DATE: 02-MAY-1995 | | | |
| CC | CC | CLASSIFICATION: 435 | | | |
| CC | CC | ATTORNEY/AGENT INFORMATION: | | | |

NAME: Meyers, Kenneth J.
 REGISTRATION NUMBER: 25,146
 REFERENCE/DOCKET NUMBER: 03495.0137-2000
 TELEPHONE: (202) 408-4000
 TELEFAX: (202) 408-4400
 INFORMATION FOR SEQ ID NO: 36:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 93 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE: 93 AA: 10002 MW: 48030 CN:
 1 5EYLILSARCVLAIVSK 93
 1 5EYLILSARCVLAIVSK 17
 Query Match 50.2% Score 114: DB 2: Length 93:
 Best Local Similarity 100.0% Pred. No. 4,56e-03:
 Matches 17: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 1A 77 5EYLILSARCVLAIVSK 93
 1 5EYLILSARCVLAIVSK 17
 RESULT 2
 ID US-09-347-594A-2 STANDARD: PRT: 2050 AA:
 XX xxxxxx
 Sequence 2, Application US/08347594A
 Sequence 2, Application: US/08347594A
 Patent No. 5849536
 GENERAL INFORMATION:
 APPLICANT: Gartinkel, Leonard
 TITLE OF INVENTION: CLONING AND PRODUCTION OF HUMAN VON
 WILLEBRAND FACTOR GPIIb BINDING DOMAIN POLYPEPTIDES AND
 METHODS OF USING SAME
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: John P. White
 STREET: 1185 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10044
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: MICROPERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: 36537-B2, Version #1.05
 FILING DATE: 05/09/94
 PRIORITY DATE: 05/09/94
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: White, John P.
 REGISTRATION NUMBER: 26,678
 REFERENCE/DOCKET NUMBER: 36537-B2
 TELEPHONE: 212-278-0400
 TELEFAX: 212-351-0525
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 2050 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE: 2050 AA: 223692 MW: 21477294 CN:

Query Match 31.7% Score 72: DB 2: Length 2050:
 Best Local Similarity 34.6% Pred. No. 2.43e-01:
 Matches 9: Conservative 6: Mismatches 9: Indels 0: Gaps 0:
 DB 198 YLILGKALSYVWDRHLSISVLKQ 223
 3 YLILSARCVLAIVSKRKVTVAIQ 28
 Query Match 31.7% Score 72: DB 2: Length 2050:
 Best Local Similarity 34.6% Pred. No. 2.43e-01:
 Matches 9: Conservative 6: Mismatches 9: Indels 0: Gaps 0:
 1A 198 YLILGKALSYVWDRHLSISVLKQ 223
 3 YLILSARCVLAIVSKRKVTVAIQ 28
 RESULT 3
 ID PC-US93-03027-2 STANDARD: PRT: 558 AA:
 XX xxxxxx
 Sequence 2, Application PC/TUS9303027
 Sequence 2, Application: PC/TUS9303027
 GENERAL INFORMATION:
 APPLICANT: LEONARD, WARREN: LEONARD,
 APPLICANT: MICHEL
 TITLE OF INVENTION: CONTROL AND/OR
 PREVENTION OF BINDING OF NF- κ B/REL:
 TITLE OF INVENTION: PREVENTION OF BINDING OF NF- κ B/REL:
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: MORGAN & FINNEGAN
 STREET: 345 PARK AVENUE
 CITY: NEW YORK
 STATE: NEW YORK
 COUNTRY: USA
 ZIP: 10154
 COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY DISK
 COMPUTER: IBM PC COMPATIBLE
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: MICROPERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PC/US93/03027
 FILING DATE: 19930401
 PRIORITY DATE: 19930401
 APPLICATION NUMBER: US/93/064,487
 FILING DATE: 06-APR-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: DOROTHY R. ADIN
 REGISTRATION NUMBER: 8-35,414
 REFERENCE/DOCKET NUMBER: 2050-101,171
 TELEPHONE: 212-758-4747
 TELEFAX: 212-758-4747
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 558
 TYPE: AMINO ACID
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 HYDROTHERMAL: NO
 ORIGINAL SOURCE:
 ORGANISM: mouse
 STRAIN:
 INDIVIDUAL ISOLATE:
 DEVELOPMENTAL STAGE:
 HAPLOTYPE:
 TISSUE TYPE:
 CELL TYPE: fibroblast
 CELL LINE: NIH-3T3
 ORGANISLE:
 FEATURE:
 NAME/KEY: RelB
 LOCATION:

| | |
|----|--|
| CC | PRIOR APPLICATION DATA: |
| CC | APPLICATION NUMBER: US 08/432,697 |
| CC | FILING DATE: 02-MAY-1995 |
| CC | CLASSIFICATION: 435 |
| CC | ATTORNEY/AGENT INFORMATION |
| CC | NAME: Meyers, Kenneth J. |
| CC | REGISTRATION NUMBER: 25,146 |
| CC | REFERENCE/DOCKET NUMBER: G3495,0137-02300 |
| CC | TELECOMMUNICATION INFORMATION: |
| CC | TELEPHONE: (202) 408-5000 |
| CC | TELEFAX: (202) 408-4400 |
| CC | INFORMATION FOR SEQ ID NO: 38: |
| CC | SEQUENCE CHARACTERISTICS: |
| CC | LENGTH: 94 amino acids |
| CC | TYPE: amino acid |
| CC | STRANDEDNESS: single |
| CC | TOPOLOGY: linear |
| CC | MOLECULE TYPE: protein |
| CC | SEQUENCE 94 AA: 10343 MW: 45288 CN: |
| CC | |
| CC | Query Match 30.8%; Score 70; DP 2; Length 94 |
| CC | Best Local Similarity 57.1%; Pred. No. 3,576; Cl: |
| CC | Matches 8; Conservative 3; Mismatches 3; Indels: |
| CC | |
| DB | 80 EYILRESILAVI 93 |
| CC | |
| QY | 2 EYILSARDLAVV 15 |
| CC | |
| CC | RESULT 5 |
| ID | US-08-619-554-4 STANDARD: PRI: 1895 AA. |
| XX | XXXXXX |
| CC | |
| DE | Sequence 4, Application US/08619554 |
| CC | Patent No. 5821353 |
| CC | GENERAL INFORMATION: |
| CC | APPLICANT: DOUGLAS, Can. Ren M. |
| CC | APPLICANT: CREBET, Gary L. |
| CC | APPLICANT: CLEMAS, Joseph |
| CC | APPLICANT: EL-SHERBEINI, Mohammed |
| CC | APPLICANT: KOHN, Forrest |
| CC | APPLICANT: KAHN, Jennifer |
| CC | APPLICANT: KELLY, Rosalind |
| CC | APPLICANT: MARRINAN, Jeanne |
| CC | APPLICANT: MORRIS, Nancy |
| CC | APPLICANT: ONISHI, Junichi |
| CC | TITLE OF INVENTION: DNA ENCODING 1,3 BETA D-GALACTAN |
| CC | NUMBER OF INVENTIONS: 8 |
| CC | CORRESPONDENCE ADDRESS: |
| CC | ADDRESSEE: JOSEPH A. COPPOLA - MERCK & CO., INC. |
| CC | STREET: 126 EAST LINCOLN AVENUE - P.O. BOX 2000 |
| CC | CITY: RAHWAY |
| CC | STATE: NJ |
| CC | COUNTRY: USA |
| CC | ZIP: 07065 |
| CC | COMPUTER READABLE FORM: |
| CC | MEDIUM TYPE: Diskette |
| CC | COMPUTER: IBM Compatib: |
| CC | OPERATING SYSTEM: DOS |
| CC | SOFTWARE: FASTSEQ for Windows Version 2.0 |
| CC | CURRENT APPLICATION DATA: |
| CC | APPLICATION NUMBER: US/08/619,554 |
| CC | FILING DATE: 01-AUG-1995 |
| CC | CLASSIFICATION: 536 |
| CC | PRIOR APPLICATION DATA: |
| CC | APPLICATION NUMBER: |
| CC | FILING DATE: |

CC ATTORNEY/AGENT INFORMATION:
CC NAME: COPPOLA, JOSEPH A.
CC REGISTRATION NUMBER: 38,413
CC REFERENCE/DOCKET NUMBER: 19104PI
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 732-594-6734
CC TELEFAX: 732-594-4720
CC TELEX:
CC INFORMATION FOR SEQ ID NO. 4:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 1895 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: unknown
CC TOPOLOGY: unknown
CC MOLECULE TYPE: protein
CC SEQUENCE 1895 AA: 216588 MW: 18982148 CN:
SQ
Query Match 30.8% Score 70: DB 2: Length 1895:
Best Local Similarity 37.5% Pred. NO. 3.57e-01:
Matches 9: Conservative 6: Mismatches 9: Indels 0: Gaps 0:
DB 658 FLLSLKGFILSTISMRCITGTY 681
QY 3 EYLISARDVLAVVSKRPMKVTVAF 26
RESULT 6
ID US-08-896-449A-2 STANDARD: PRI: 2813 AA.
XX
AC xxxxxx
DE
XX
XX
XX
Sequence 2, Application US/08896449A
Patent No. 640143
GENERAL INFORMATION:
APPLICANT: Vecta, Patrick J.
APPLICANT: Yuzbysyan-Gurkay, Vilma
APPLICANT: Schall, William D.
APPLICANT: Brewer, George J.
TITLE OF INVENTION: DNA ENCODING CANINE VON WILLEBRAND
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hainnes, Dickey & Pierce, P.C.
STREET: 5445 Corporate Drive
CITY: Troy
STATE: Michigan
COUNTRY: USA
ZIP: 48062
COMPUTER READABLE FORM
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/896.449A
FILING DATE: 18-JUL-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Smith, Deann F.
REFERENCE/DOCKET NUMBER: 2115-001226
TELECOMMUNICATION INFORMATION:
TELEPHONE: 248-641-1600
TELEFAX: 248-641-0270
TELEX: 287637
INFORMATION FOR SEQ ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2813 amino acids
TYPE: amino acid
TOPOLOGY: linear

CC MOLECULE TYPE: protein
CC SEQUENCE 2813 AA: 309523 : 40158915 CN:
SQ
Query Match 30.8% Score 70: DB 3: Length 2813:
Best Local Similarity 29.8% Pred. NO. 3.57e-01:
Matches 8: Conservative 11: Mismatches 9: Indels 0: Gaps 0:
DB 960 QYVILLGNALSVWDRSLSVTLKR 986
QY 2 EYLISARDVLAVVSKRPMKVTVAF 28
RESULT 7
ID US-09-144-925-23 STANDARD: PRI: 254 AA.
XX
AC xxxxxx
DE
XX
XX
XX
Sequence 23, Application US/09144925
Patent No. 5951979
GENERAL INFORMATION:
APPLICANT: Toks, Nicholas
APPLICANT: Flint, Andrew J.
TITLE OF INVENTION: SUBSTRATE TRAPPING PROTEIN
TITLE OF INVENTION: TYROSINE PHOSPHATASES
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: HAMILTON, B. K. SMITH & REYNOLDS, P.C.
STREET: Two Militia Dr.
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02421-4799
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows 95
SOFTWARE: FastSeq for Windows Version 2.00
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/144.925
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/85.992
FILING DATE: July 25, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Sheridan, Patricia
REGISTRATION NUMBER: 38427
REFERENCE/DOCKET NUMBER: 2115-001226
TELECOMMUNICATION INFORMATION:
TELEPHONE: 781-861-0240
TELEFAX: 781-861-9540
TELEX:
INFORMATION FOR SEQ ID NO. 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 254 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE 254 AA: 26034 MW: 364632 CN:
SQ
Query Match 30.0% Score 68: DB 2: Length 254:
Best Local Similarity 46.7% Pred. NO. 5.21e-01:
Matches 7: Conservative 4: Mismatches 4: Indels 0: Gaps 0:
DB 230 IQRKMKMVTQTQF 244
QY 15 VSKRPMKVTVAFNQF 29

```

RESULT 8
ID US-08-685-592-23 STANDARD: PRT: 254 AA.
XX AC xxxxxx
XX DT
XX DE
XX SEQUENCE 23, Application US/08685992
XX
XX SEQUENCE 23, Application US/08685992
XX Patent No. 5921238
XX GENERAL INFORMATION:
XX APPLICANT: IONAS, Nicholas
XX APPLICANT: FILL, Andrew J.
XX TITLE OF INVENTION: SUBSTRATE TRAPPING PROTEIN
XX TITLE OF INVENTION: TYROSINE PHOSPHATASES
XX NUMBER OF SEQUENCES: 35
XX CORRESPONDENCE ADDRESS:
XX ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
XX STREET: Two Militia Drive
XX CITY: Lexington
XX STATE: MA
XX COUNTRY: USA
XX ZIP: 02173
XX COMPUTER READABLE FORM:
XX MEDIUM TYPE: Diskette
XX COMPUTER: IBM Compatible
XX OPERATING SYSTEM: Windows 95
XX SOFTWARE: FastSeq for Windows Version 2.0b
XX CURRENT APPLICATION DATA:
XX APPLICATION NUMBER: US/08/685,992
XX FILING DATE: 25-JUL-1996
XX CLASSIFICATION: 435
XX PRIOR APPLICATION DATA:
XX APPLICATION NUMBER:
XX FILING DATE:
XX ATTORNEY/AGENT INFORMATION:
XX NAME: Granahan, Patricia
XX REGISTRATION NUMBER: 32,227
XX REFERENCE/DOCKET NUMBER: CSHL96-03
XX TELECOMMUNICATION INFORMATION:
XX TELEPHONE: 781-861-6240
XX TELEFAX: 781-861-9540
XX TELEX:
XX INFORMATION FOR SEQ ID NO: 23:
XX SEQUENCE CHARACTERISTICS:
XX LENGTH: 254 amino acids
XX TYPE: amino acid
XX STRANDEDNESS: single
XX TOPOLOGY: linear
XX MOLECULE TYPE: peptide
XX SEQUENCE 254 AA: 25254 MW: 364832 CN.

Query Match 30.0% Score 68: DB 2: Length 254:
Best Local Similarity 46.7% Pref. No. 5,21e-01:
Matches 7: Conservative 4: Mismatches 4: Indels 0: Gaps 0:

DB 23C IRKQRMKVKVQTFQF 244
: : : : : : : : : :
QY 15 YSKRRMKVVAENQF 29

RESULT 9
ID US-07-857-224B-79 STANDARD: PRT: 269 AA.
XX AC xxxxxx
XX DT
XX DE
XX SEQUENCE 79, Application US/07857224B
XX
XX SEQUENCE 79, Application US/07857224B
XX Patent No. 592428
XX GENERAL INFORMATION:
XX APPLICANT: Moosa Mohammed, Joseph Schlessinger,
XX APPLICANT: and Stevan R. Hubbard
XX TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN
XX TITLE OF INVENTION: OF NCN-INSULIN RECEPTOR TYROSINE KINASE

Patent No. 5958784
GENERAL INFORMATION:
APPLICANT: Benner, Steven A.
TITLE OF INVENTION: Predicting Folded Structures of Proteins
NUMBER OF SEQUENCES: 114
CORRESPONDENCE ADDRESS:
ADDRESSEE: Steven A. Benner
STREET: Hadlaubstrasse 151
CITY: Zurich
STATE: none
COUNTRY: Switzerland
ZIP: (note: this is an international post code) CH-8092
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage
COMPUTER: Apple Macintosh
OPERATING SYSTEM: Macintosh 7.0
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US 7857,224B
FILING DATE: 03/25/92
CLASSIFICATION: 436
PRIOR APPLICATION DATA: none
TELECOMMUNICATION INFORMATION:
TELEPHONE: (International) 41 1 632 2830
TELEFAX: (International) 41 1 262 2437
TELEX: none
INFORMATION FOR SEQ ID NO: 79:
SEQUENCE CHARACTERISTICS:
LENGTH: 269
TYPE: amino acid
MOLECULE TYPE: linear
DESCRIPTION: protein
ORIGINAL SOURCE:
ORGANISM: human
FEATURE: Protein kinase; Table 8 Column 90
PUBLICATION INFORMATION:
AUTHORS:
AUTHORS: Hanks, S. K.
AUTHORS: Quinn, A. M.
AUTHORS: Hunter, T.
TITLE: The protein kinase family
JOURNAL: Science
VOLUME: 241
PAGES: 42-52
DATE: 1988
SEQUENCE 269 AA: 30804 MW: 43458 CN.

Query Match 30.0% Score 68: DB 2: Length 269:
Best Local Similarity 44.4% Pref. No. 5,21e-01:
Matches 8: Conservative 3: Mismatches 3: Indels 1: Gaps 1:

DB 133 LAARNILVAESPCKMKIS 149
: : : : : : : : : :
QY 6 LSARDVLAVSKRRMKVT 23

RESULT 10
ID US-08-701-191A-27 STANDARD: PRT: 304 AA.
XX AC xxxxxx
XX DT
XX DE
XX SEQUENCE 27, Application US/08701191A
XX
XX SEQUENCE 27, Application US/08701191A
XX Patent No. 5942428
XX GENERAL INFORMATION:
XX APPLICANT: Moosa Mohammed, Joseph Schlessinger,
XX APPLICANT: and Stevan R. Hubbard
XX TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN
XX TITLE OF INVENTION: OF NCN-INSULIN RECEPTOR TYROSINE KINASE
```

CC NUMBER OF SEQUENCES: 41
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Lyon & Lyon
 CC STREET: 613 West Fifth Street
 CC CITY: Suite 4700
 CC STATE: Los Angeles
 CC COUNTRY: California
 CC ZIP: 90071-2066
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
 CC OPERATING SYSTEM: IBM Compatible
 CC SOFTWARE: FASTSEC for Windows 2.0
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/701.191A
 CC FILING DATE: August 21, 1996
 CC CLASSIFICATION: 59C
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER:
 CC FILING DATE:
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Warburg, Richard J.
 CC REGISTRATION NUMBER: 32,327
 CC REFERENCE/DOCKET NUMBER: 227/488
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (213) 489-1600
 CC TELEFAX: (213) 905-0440
 CC TELEX: 67-3510
 CC INFORMATION FOR SEQ ID NO: 27:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 304 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS: single
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC IMMEDIATE SOURCE:
 CC CLONE: Ret
 CC SEQUENCE 304 AA: 34556 MW: 480234 CN:

Query Match 30.0% Score 68: DB 2: Length 304:
 Best Local Similarity 44.4% Pred. No. 5.21e-01:
 Matches 8: Conservative 6: Mismatches 3: Indels 1: Gaps 1:

DB 165 LAARNIL-VAEGRKMKIS 161
 QY 6 LSARDVLAVVSKRMRKVT 23

RESULT 11
 ID US-08-548-057A-15 STANDARD: 160 316 AA
 AC xxxxxx
 DT
 XX

Sequence 15: Application US/08638957A
 Patent No. 5998157

GENERAL INFORMATION:
 APPLICANT: Breitman, Martin L.
 APPLICANT: Rossant, Janet
 APPLICANT: Duront, Daniel J.
 APPLICANT: Yamaguchi, Terry P.
 TITLE OF INVENTION: NO. 5998187e1 Receptor Tyrosine Kinase
 NUMBER OF SEQUENCES: 32
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Bereskin & Parr
 STREET: 40 King Street West
 CITY: Toronto
 STATE: Ontario
 COUNTRY: Canada
 ZIP: M5H 3Y2

CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC OPERATING SYSTEM: IBM PC compatible
 CC SOFTWARE: PatentIn Release #1.0, Version #1.30
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/838.957A
 CC FILING DATE: 23-APR-1997
 CC CLASSIFICATION: 435
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Kurdydk, Linda M.
 CC REGISTRATION NUMBER: 34,971
 CC REFERENCE/DOCKET NUMBER: 3153-212
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (416) 364-7311
 CC TELEFAX: (416) 361-1398
 CC INFORMATION FOR SEQ ID NO: 15:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 316 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS: single
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC IMMEDIATE SOURCE:
 CC CLONE: Ret
 CC SEQUENCE 316 AA: 35596 MW: 1-0817 CN:
 CC
 CC Query Match 30.0% Score 68: DB 2: Length 316:
 CC Best Local Similarity 44.4% Pred. No. 5.21e-01:
 CC Matches 8: Conservative 6: Mismatches 3: Indels 1: Gaps 1:
 CC
 CC DB 152 LAARNIL-VAEGRKMKIS 168
 CC QY 6 LSARDVLAVVSKRMRKVT 23
 CC
 CC RESULT 12
 CC ID US-08-278-089A-16 STANDARD: 167 316 AA
 CC AC xxxxxx
 CC DT
 CC XX
 CC
 CC Sequence 16: Application US/08278089A
 CC Patent No. 56817-4
 CC GENERAL INFORMATION:
 CC APPLICANT: Breitman, Martin L.
 CC APPLICANT: Rossant, Janet
 CC APPLICANT: Yamaguchi, Terry P.
 CC TITLE OF INVENTION: NO. 56817-4 Receptor Tyrosine Kinase
 CC NUMBER OF SEQUENCES: 34
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Bereskin & Parr
 CC STREET: 40 King Street West
 CC CITY: Toronto
 CC STATE: Ontario
 CC COUNTRY: Canada
 CC ZIP: M5H 3Y2
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC OPERATING SYSTEM: IBM PC compatible
 CC SOFTWARE: PatentIn Release #1.0, Version #1.30
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/278.089A
 CC FILING DATE: 20-JUL-1994
 CC CLASSIFICATION: 530
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Kurdydk, Linda M.
 CC REGISTRATION NUMBER: 34,971


```
CC . REFERENCE/DOCKET NUMBER: 3153-111
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (416) 364-7311
CC TELEFAX: (416) 361-1398
CC INFORMATION FOR SEQ ID NO: 16:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 316 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC IMMEDIATE SOURCE:
CC CLONE: Ret
CC SEQUENCE 316 AA: 35696 MW: 52081.7 CN:

Query Match 30.0% Score 67: DB 1: Length 316:
Best Local Similarity 44.4% Pred. No. 5.21e-01:
Matches 8: Conservative 5: Mismatches 3: Indels 1: Gaps 1:

DB 152 LAARN:L-VAEGRKMKIS 168
QY 6 LSARDVLAVVSKRMKRV 23

RESULT 13
ID US-09-824-878-3 STANDARD: PRT: 368 AA.
XX
AC xxxxxx
XX
XX
DE
XX
XX
Sequence 3: Application US/08824878
Patent No. 5981221
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
TITLE OF INVENTION: NOVEL HISTONE FUSION PROT
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Inocyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/824,878
FILING DATE: Filed Herewith.
CLASSIFICATION: 425
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-C255 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 368 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:

CC . REFERENCE/DOCKET NUMBER: 3153-111
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (416) 364-7311
CC TELEFAX: (416) 361-1398
CC INFORMATION FOR SEQ ID NO: 16:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 316 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: cDNA
CC IMMEDIATE SOURCE:
CC CLONE: Ret
CC SEQUENCE 316 AA: 35696 MW: 52081.7 CN:

Query Match 29.5% Score 67: DB 2: Length 368:
Best Local Similarity 32.3% Pred. No. 6.30e-01:
Matches 10: Conservative 10: Mismatches 10: Indels 1: Gaps 1:

DB 283 EELLEKTVKNCALADERKLLK-SIAFFSIGS 312
QY 1 EYLILSARDVLAVVSKRMKRVVAVAGFGP 31

RESULT 14
ID US-08-119-361-5 STANDARD: PRT: 737 AA.
XX
AC xxxxxx
XX
XX
DE
XX
XX
Sequence 5: Application US/08119361
Patent No. 5523390
GENERAL INFORMATION:
APPLICANT: Travis, James
APPLICANT: Potempa, Jan
APPLICANT: Barr, Philip J.
APPLICANT: Pavloff, Nadine
TITLE OF INVENTION: Porphyromonas gingivalis Arginine-specific Proteinase
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee and Winner, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: CO
COUNTRY: USA
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/119,361
FILING DATE: 10-SEP-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Peibert, Donna M.
REGISTRATION NUMBER: 4474
REFERENCE/DOCKET NUMBER: 21111
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303-439-8086
TELEFAX: 303-499-8089
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 737 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE 737 AA: 81332 MW: 1874143 CN:

Query Match 29.5% Score 67: DB 1: Length 737:
Best Local Similarity 52.0% Pred. No. 6.30e-01:
Matches 13: Conservative 9: Mismatches 9: Indels 2: Gaps 2:

DB 87 ILS-RS-LAVSDTREMKVEVYSSKF 109
QY 5 ILSARDVLAVVSKRMKRVTVAFNQ 29

RESULT 15
ID US-08-336-308A-4 STANDARD: PRT: 737 AA.
```


1








CC CLASSIFICATION: 435
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Meyers, Kenneth J.
 CC REGISTRATION NUMBER: 25,146
 CC REFERENCE/DOCKET NUMBER: 03495.0137-02000
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (202) 408-4000
 CC TELEFAX: (202) 408-4400
 CC INFORMATION FOR SEQ ID NO: 36:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 93 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS: single
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC SEQUENCE 93 AA: 10002 MW: 48030 CN:
 SV
 Query Match 50.2% Score 114: DB 10: length 93:
 Best Local Similarity 100.0% Pred. No. 1.68e-02:
 Matches 17: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 DE 77 EYLLSARDVLAVVSK 93
 QY I EYLLSARDVLAVVSK 17
 CC
 RESULT 2
 CC US-09-142-697-45 STANDARD: PRT: 93 AA:
 CC XXXXX
 CC
 CC Sequence 36, Application US/08432697
 CC Sequence 36, Application US/08432697
 CC GENERAL INFORMATION:
 CC APPLICANT: Labigne, Agnes
 CC APPLICANT: Sauerbaum, Sebastien
 CC APPLICANT: Ferrero, Richard L.
 CC APPLICANT: Philbert, Jean-Michel
 CC TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST
 CC TITLE OF INVENTION: HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE
 CC TITLE OF INVENTION: COMPOSITIONS, AND NUCLEIC ACID SEQUENCES ENCODING SAID
 CC TITLE OF INVENTION: POLYPEPTIDES
 CC NUMBER OF SEQUENCES: 44
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Finneqan, Henderson, Farabow, Garrett &
 CC ADDRESSEE: Cannon
 CC STREET: 1200 I Street, N.W.
 CC CITY: Washington
 CC STATE: D.C.
 CC COUNTRY: USA
 CC ZIP: 20005-3315
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patent Release #1.0, Version #1.30
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/432,697
 CC FILING DATE: 02-MAY-1995
 CC CLASSIFICATION: 424
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Meyers, Kenneth J.
 CC REGISTRATION NUMBER: 25,146
 CC REFERENCE/DOCKET NUMBER: 03495.0137-00000
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (202) 408-4000
 CC TELEFAX: (202) 408-4400
 CC INFORMATION FOR SEQ ID NO: 36:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 93 amino acids

CC TYPE: amino acid
 CC STRANDEDNESS: single
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC SEQUENCE 93 AA: 10002 MW: 48030 CN:
 SV
 Query Match 50.2% Score 114: DB 9: length 93:
 Best Local Similarity 100.0% Pred. No. 1.68e-02:
 Matches 17: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 DE 77 EYLLSARDVLAVVSK 93
 QY I EYLLSARDVLAVVSK 17
 CC
 RESULT 3
 CC US-09-146-940-236 STANDARD: PRT: 19 AA:
 CC XXXXX
 CC
 CC Sequence 236, Application US/08468940
 CC Sequence 236, Application US/08468940
 CC GENERAL INFORMATION:
 CC APPLICANT: Griffith, Irvin C.
 CC APPLICANT: Pollock, Joanne
 CC APPLICANT: Bond, Julian F.
 CC APPLICANT: Garman, Richard D.
 CC APPLICANT: Kuo, Mei-Chang
 CC APPLICANT: Yeung, Siu-mei H.
 CC APPLICANT: Brauer, Andrew
 CC APPLICANT: Exley, Mark A.
 CC APPLICANT: Powels, Steven P.
 CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
 CC TITLE OF INVENTION: Japanese Cedar Folles
 CC NUMBER OF SEQUENCES: 261
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
 CC STREET: 610 Lincoln St.
 CC CITY: Waltham
 CC STATE: MA
 CC COUNTRY: USA
 CC ZIP: 02154
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patent Release #1.0, Version #1.2
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/468,940
 CC FILING DATE: 08/07/98
 CC CLASSIFICATION: 424
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/350,225
 CC FILING DATE: December 6, 1994
 CC APPLICATION NUMBER: 08/226,248
 CC FILING DATE: April 8, 1994
 CC APPLICATION NUMBER: 07/339,990
 CC FILING DATE: September 1, 1992
 CC APPLICATION NUMBER: PCT/US93/00139
 CC FILING DATE: January 15, 1993
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Darlene A. Vanstone
 CC REGISTRATION NUMBER: 35,729
 CC REFERENCE/DOCKET NUM: 025.6 US (IMI-0280P2)
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (617) 466-6000
 CC TELEFAX: (617) 466-6040
 CC INFORMATION FOR SEQ ID NO: 236:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 19 amino acids

CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: peptide
 CC FRAGMENT TYPE: internal
 SQ SEQUENCE 19 AA: 218 MW: 2058 CN:

Query Match 46.3% Score 105: DB 10: Length 19;
 Best Local Similarity 87.5%: Pred. No. 1.39e-01;
 Matches 14: Conservative 1: Mismatches 1: Indels 0: Gaps 0:

DB 2 KSKMKVTVAFNQFGPN 17
 I: |||||
 QY 17 KRRMKVTVAFNQFGPN 32

RESULT 4
 ID US-09-350-225-236 STANDARD: PRT: 19 AA:

XX xxxxxx

Sequence 236, Application US/09350225

Sequence 236, Application US/09350225

GENERAL INFORMATION:

APPLICANT: Griffith, Irwin J.

APPLICANT: Pollock, Joanne

APPLICANT: Bond, Julian F.

APPLICANT: Garman, Richard D.

APPLICANT: Kuo, Mei-Chang

APPLICANT: Yeung, Siu-wei H.

APPLICANT: Brauer, Andrew

APPLICANT: Exley, Mark A.

APPLICANT: Powers, Steven P.

TITLE OF INVENTION: Allergenic Proteins And Peptides From

NUMBER OF SEQUENCES: 261

CORRESPONDENCE ADDRESS:

ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.

STREET: 610 Lincoln St

CITY: Waltham

STATE: MA

COUNTRY: USA

ZIP: 02154

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/93/350,225

FILING DATE: December 6, 1994

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/226,248

FILING DATE: April 8, 1994

APPLICATION NUMBER: 07/938,990

FILING DATE: September 1, 1992

APPLICATION NUMBER: PCT/US93/00139

FILING DATE: January 15, 1993

ATTORNEY/AGENT INFORMATION:

NAME: Darlene A. Vanstone

REGISTRATION NUMBER: 35,729

REFERENCE/DOCKET NUMBER: 025.6 US (IMI-028CP2)

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 466-6000

TELEFAX: (617) 466-6040

INFORMATION FOR SEQ ID NO: 236:

SEQUENCE CHARACTERISTICS:

LENGTH: 19 amino acids

TYPE: amino acid

TOPOLOGY: linear

CC MOLECULE TYPE: peptide
 CC FRAGMENT TYPE: internal
 SQ SEQUENCE 19 AA: 2181 MW: 2058 CN:

Query Match 46.3% Score 105: DB 8: Length 19;
 Best Local Similarity 87.5%: Pred. No. 1.39e-01;
 Matches 14: Conservative 1: Mismatches 1: Indels 0: Gaps 0:

DB 2 KSKMKVTVAFNQFGPN 17
 I: |||||
 QY 17 KRRMKVTVAFNQFGPN 32

RESULT 5
 ID US-08-467-006-236 STANDARD: PRT: 19 AA:

XX xxxxxx

Sequence 236, Application US/08467006

Sequence 236, Application US/08467006

GENERAL INFORMATION:

APPLICANT: Griffith, Irwin J.

APPLICANT: Pollock, Joanne

APPLICANT: Bond, Julian F.

APPLICANT: Garman, Richard D.

APPLICANT: Kuo, Mei-Chang

APPLICANT: Yeung, Siu-wei H.

APPLICANT: Brauer, Andrew

APPLICANT: Exley, Mark A.

APPLICANT: Powers, Steven P.

TITLE OF INVENTION: Allergenic Proteins And Peptides From

NUMBER OF SEQUENCES: 261

CORRESPONDENCE ADDRESS:

ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.

STREET: 610 Lincoln St

CITY: Waltham

STATE: MA

COUNTRY: USA

ZIP: 02154

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/98/467,006

FILING DATE: June 4, 1998

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/350,225

FILING DATE: December 6, 1994

ATTORNEY/AGENT INFORMATION:

NAME: Jane E. Remillard

REGISTRATION NUMBER: 38,872

REFERENCE/DOCKET NUMBER: 025.6 USD6 (IMI-028CPD6)

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400

TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 236:

SEQUENCE CHARACTERISTICS:

LENGTH: 19 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FRAGMENT TYPE: internal

SEQUENCE 19 AA: 2181 MW: 2058 CN:

Query Match 46.3% Score 105: DB 10: Length 19;
 Best Local Similarity 87.5%: Pred. No. 1.39e-01;

Matches 14: Conservative 1: Mismatches 1: Indels 0: Gaps 0:

Db 2 KSKMKVTVAFNQFGPN 17
1: |||||.....
Qy 17 KRMKVTVAFNQFGPN 32

RESULT 6

ID US-08-467-023-236 STANDARD: PRT: 19 AA.

XX xxxxxx

Sequence 236, Application US/08467023

Sequence 236, Application US/08467023

GENERAL INFORMATION:
APPLICANT: Griffith, Irwin J.
APPLICANT: Pollock, Joanne
APPLICANT: Bond, Julian F.
APPLICANT: Garman, Richard D.
APPLICANT: Kuo, Mei-Chang
APPLICANT: Yeung, Siu-mei H.
APPLICANT: Brauer, Andrew
APPLICANT: Exley, Mark A.
APPLICANT: Powers, Steven P.
TITLE OF INVENTION: Allergenic Proteins And Peptides From
TITLE OF INVENTION: Japanese Cedar Pollen
NUMBER OF SEQUENCES: 261
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
STREET: 610 Lincoln St
CITY: Waltham
STATE: MA
COUNTRY: USA
ZIP: 02154

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,023
FILING DATE: June 6, 1995
CLASSIFICATION: 424

REGISTRATION NUMBER: 38,872

REFERENCE/DOCKET NUMBER: 925,6 (US02 (IM-018GPD2)

TELEPHONE: (617) 227-7400

TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 236:

SEQUENCE CHARACTERISTICS:

LENGTH: 19 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FRAGMENT TYPE: internal

SEQUENCE 19 AA: 2181 MW: 2058 CN:

Query Match 45.3% Score 105; DB 10: Length 19;

Best Local Similarity 87.5% Pred No. 1.39e-01;

Matches 14: Conservative 1: Mismatches 1: Indels 0: Gaps 0:

Db 2 KSKMKVTVAFNQFGPN 17

1: |||||.....

Qy 17 KRMKVTVAFNQFGPN 32

RESULT 7

ID US-08-467-697-236 STANDARD: PRT: 19 AA.

XX xxxxxx

Sequence 236, Application US/08467697

Sequence 236, Application US/08467697

GENERAL INFORMATION:
APPLICANT: Griffith, Irwin J.
APPLICANT: Pollock, Joanne
APPLICANT: Bond, Julian F.
APPLICANT: Garman, Richard D.
APPLICANT: Kuo, Mei-Chang
APPLICANT: Yeung, Siu-mei H.
APPLICANT: Brauer, Andrew
APPLICANT: Exley, Mark A.
APPLICANT: Powers, Steven P.
TITLE OF INVENTION: Allergenic Proteins And Peptides From
TITLE OF INVENTION: Japanese Cedar Pollen
NUMBER OF SEQUENCES: 261
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
STREET: 610 Lincoln St
CITY: Waltham
STATE: MA
COUNTRY: USA
ZIP: 02154

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,697
FILING DATE: June 6, 1995
CLASSIFICATION: 424

REGISTRATION NUMBER: 38,872

REFERENCE/DOCKET NUMBER: 925,6 (US02 (IM-018GPD2)

TELEPHONE: (617) 227-7400

TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 236:

SEQUENCE CHARACTERISTICS:

LENGTH: 19 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FRAGMENT TYPE: internal

SEQUENCE 19 AA: 2181 MW: 2058 CN:

Query Match 45.3% Score 105; DB 10: Length 19;

Best Local Similarity 87.5% Pred No. 1.39e-01;

Matches 14: Conservative 1: Mismatches 1: Indels 0: Gaps 0:

Db 2 KSKMKVTVAFNQFGPN 17

1: |||||.....

Qy 17 KRMKVTVAFNQFGPN 32

RESULT 8

ID US-08-468-940-222 STANDARD: PRT: 29 AA.

XX xxxxxx

AC xxxxxx

XX

DT XX
DE XX
XX
Sequence 222, Application US/09458940
Sequence 222, Application US/09458940
GENERAL INFORMATION:
APPLICANT: Griffith, Irwin J.;
APPLICANT: Pollock, Joanne;
APPLICANT: Bond, Julian F.;
APPLICANT: Garman, Richard D;
APPLICANT: Kuo, Mei-Chang;
APPLICANT: Yeung, Siu-mei H.;
APPLICANT: Brauer, Andrew;
APPLICANT: Exley, Mark A.;
APPLICANT: Powers, Steven P.
TITLE OF INVENTION: Allergenic Proteins And Peptides From
TITLE OF INVENTION: Japanese Cedar Pollen
NUMBER OF SEQUENCES: 26;
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
STREET: 610 Lincoln St
CITY: Waltham
STATE: MA
COUNTRY: USA
ZIP: 02154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468-440
FILING DATE: December 6, 1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/350,225
FILING DATE: December 6, 1994
APPLICATION NUMBER: 08/226,248
FILING DATE: April 8, 1994
APPLICATION NUMBER: 07/938,990
FILING DATE: September 1, 1992
APPLICATION NUMBER: PCT/US93/00139
FILING DATE: January 15, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Darlene A. Vanstone
REGISTRATION NUMBER: 35,729
REFERENCE/DOCKET NUMBER: 025.6 US (IM1-028CP2)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 466-6000
TELEFAX: (617) 466-6040
INFORMATION FOR SEQ ID NO: 222
SEQUENCE CHARACTERISTICS:
LENGTH: 29 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
SEQUENCE 29 AA: 3313 MW: 4103 CN:
Query Match 46.3%; Score 105; DB 10; Length 29;
Best Local Similarity 87.5%; Pred. NO. 1.39e-01;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Db 1 KSKMKVTVAFNFGPN 16
QY 17 KRMKVTVAFNFGPN 32
RESULT 9
ID US-08-350-225-222 STANDARD: PRT: 29 AA.
XX
AC xxxxxx
XX
DT XX

DT XX
DE XX
XX
Sequence 222, Application US/08350225
Sequence 222, Application US/08350225
GENERAL INFORMATION:
APPLICANT: Griffith, Irwin J.;
APPLICANT: Pollock, Joanne;
APPLICANT: Bond, Julian F.;
APPLICANT: Garman, Richard D;
APPLICANT: Kuo, Mei-Chang;
APPLICANT: Yeung, Siu-mei H.;
APPLICANT: Brauer, Andrew;
APPLICANT: Exley, Mark A.;
APPLICANT: Powers, Steven P.
TITLE OF INVENTION: Allergenic Proteins And Peptides From
TITLE OF INVENTION: Japanese Cedar Pollen
NUMBER OF SEQUENCES: 26;
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
STREET: 610 Lincoln St
CITY: Waltham
STATE: MA
COUNTRY: USA
ZIP: 02154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/350,225
FILING DATE: December 6, 1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/226,248
FILING DATE: April 8, 1994
APPLICATION NUMBER: 07/938,990
FILING DATE: September 1, 1992
APPLICATION NUMBER: PCT/US93/00139
FILING DATE: January 15, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Darlene A. Vanstone
REGISTRATION NUMBER: 35,729
REFERENCE/DOCKET NUMBER: 025.6 US (IM1-028CP2)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 466-6000
TELEFAX: (617) 466-6040
INFORMATION FOR SEQ ID NO: 222
SEQUENCE CHARACTERISTICS:
LENGTH: 29 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
SEQUENCE 29 AA: 3313 MW: 4103 CN:
Query Match 46.3%; Score 105; DB 8; Length 29;
Best Local Similarity 87.5%; Pred. NO. 1.39e-01;
Matches 14; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Db 1 KSKMKVTVAFNFGPN 16
QY 17 KRMKVTVAFNFGPN 32
RESULT 10
ID US-08-467-023-222 STANDARD: PRT: 29 AA.
XX
AC xxxxxx
XX
DT XX


```
CC STREET: 610 Lincoln St  
CC CITY: Waltham  
CC STATE: MA  
CC COUNTRY: USA  
CC ZIP: 02154  
CC COMPUTER READABLE FORM:  
CC MEDIUM TYPE: Floppy disk  
CC COMPUTER: IBM PC compatible  
CC OPERATING SYSTEM: PC-DOS/VMS-DOS  
CC SOFTWARE: Patent Release #1.0, Version #1.25  
CC CURRENT APPLICATION DATA:  
CC APPLICATION NUMBER: US/08/350,225  
CC FILING DATE: December 6, 1994  
CC CLASSIFICATION: 424  
CC PRIORITY APPLICATION DATA:  
CC APPLICATION NUMBER: 08/226,246  
CC FILING DATE: April 8, 1994  
CC APPLICATION NUMBER: 07/938,990  
CC FILING DATE: September 1, 1992  
CC APPLICATION NUMBER: PCI/US93/00139  
CC FILING DATE: January 15, 1993  
CC ATTORNEY/AGENT INFORMATION:  
CC NAME: Darlene A. Vanstone  
CC REGISTRATION NUMBER: 35,729  
CC REFERENCE/DOCKET NUMBER: 025.6 US (MI-028CP2)  
CC TELECOMMUNICATION INFORMATION:  
CC TELEPHONE: (617) 466-8000  
CC TELEFAX: (617) 466-6040  
CC SEQUENCE CHARACTERISTICS:  
CC LENGTH: 26 amino acids  
CC TYPE: amino acid  
CC TOPOLOGY: linear  
CC MOLECULE TYPE: peptide  
CC FRAGMENT TYPE: internal  
CC  
SQ SEQUENCE 26 AA; 2956 MW; 3131 CN;  
  
Query Match 45.4%; Store 103; DB 8; Length 26;  
Best Local Similarity 93.3%; Pr.d.No. 2.21e-01;  
Matches 14; Conservative 3; Mismatches 1; Indels 0;  
  
Db I RSMKVTVAFGNGFN 15  
I | | | | | | |  
QY 18 PRMKVTVAFGNGFN 32  
  
RESULT 14  
ID US-08-350-225-229 STANDARD; PRI: 26 AA.  
XX xxxxxx  
AC  
XX  
D1 XX  
DE XX  
XX  
Sequence 225, Application: US/08350225  
  
Sequence 229, Application US/0850225  
GENERAL INFORMATION:  
APPLICANT: Griffith, Irwin J.;  
APPLICANT: Pollock, Joanne;  
APPLICANT: Bond, Julian F.;  
APPLICANT: Garman, Richard D.;  
APPLICANT: Kuo, Mei-Chang;  
APPLICANT: Yeung, Siu-mei H.;  
APPLICANT: Brauer, Andrew;  
APPLICANT: Exley, Mark A.;  
APPLICANT: Powers, Steven P.  
TITLE OF INVENTION: Allergenic Proteins And Peptides From  
NUMBER OF SEQUENCES: 261  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.,  
CITY: Waltham,  
STREET: 610 Lincoln St
```

```

CC ZIP: 02154
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/68/467,023
CC FILING DATE: June 6, 1996
CC CLASSIFICATION: 424
CC PRIORITY APPLICATION DATA:
CC APPLICATION NUMBER: 08/150,226
CC FILING DATE: December 1, 1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Jane E. Remillard
CC REGISTRATION NUMBER: 56,72
CC REFERENCE/DISCLET NUMBER: 45,1 USA (IMP-02RPLP2)
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 227-7
CC TELEFAX: (617) 227-594
CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 26 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC FRAGMENT TYPE: Internal
CC QUERY MATCH: 45.4%; Score 10; DB 10; Length 29;
CC Best Local Similarity 93.3%; Pred.No.2 2le-01;
CC Matches 14; Conservative 0; Mismatches 1; Indels
Cc 1 RSMKVTVAFTNGFGPN 15
Cq 18 RSMKVTVAFTNGFGPN 32
Search completed: Mon Jun 19 16:26:53 2000
Job time : 19 secs.
```

[illegible]

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

```

updatechpp protein - protein database search, using Smith-Waterman algorithm
Run on:      Mon Jun 19 16:12:33 2000:  Maspar time 6.47 Seconds
              233.455 Mill. on cell updates/sec
              tabular output not generated.

```

```

>US-09-142-524A-10
>Description: (1-32) from US09142524A.pcp
>Perfect Score: 227
>Sequence: 1: ERYALSAROV LAVYSKRMRKVTVAFNCGPN 32

```

Scoring table: PAM 150
Gap 11

Searched: 142080 seqs, 47172436 residues

Post-processing: Minimum Match 08
Listing first 45 summaries

```
Database: pir52
1:pir1 2:pir2 3:pir3 4:pir4
```

Statistics: Mean: 34.452; Variance: 60.396; scale: 0.570

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

| Result
No. | Score | Query
Match | Length | DB ID | Description | Filed No. |
|---------------|-------|----------------|--------|-------|-------------|------------|
| 1 | 114 | 50.2 | 100 | 4 | S25180 | 1 172e-06 |
| 2 | 114 | 50.2 | 100 | 2 | JHQ556 | 1 172e-06 |
| 3 | 114 | 50.2 | 100 | 1 | BVMYBA | 1 172e-06 |
| 4 | 114 | 50.2 | 169 | 2 | S72818 | 1 172e-06 |
| 5 | 105 | 46.3 | 102 | 2 | S37365 | 1 172e-06 |
| 6 | 103 | 45.4 | 102 | 2 | A41325 | 1 143e-04 |
| 7 | 98 | 43.2 | 374 | 2 | JC2124 | 1 100e-03 |
| 8 | 98 | 43.2 | 374 | 2 | JC2123 | 1 100e-03 |
| 9 | 88 | 38.8 | 242 | 2 | E71621 | 1 4.27e-02 |
| 10 | 84 | 37.0 | 455 | 2 | T0C856 | 1 1.80e-01 |
| 11 | 80 | 35.2 | 103 | 1 | BVWGGS | 1 7.31e-01 |
| 12 | 80 | 35.2 | 103 | 2 | A36721 | 1 7.31e-01 |
| 13 | 80 | 35.2 | 310 | 2 | T11320 | 1 7.31e-01 |
| 14 | 80 | 35.2 | 328 | 2 | T11938 | 1 7.31e-01 |
| 15 | 79 | 34.8 | 100 | 1 | BVMY7B | 1 1.03e+00 |
| 16 | 79 | 34.8 | 869 | 1 | A47257 | 1 1.03e+00 |
| 17 | 78 | 34.4 | 103 | 2 | T06830 | 1 1.45e+00 |
| 18 | 78 | 34.4 | 230 | 1 | P53610 | 1 1.45e+00 |
| 19 | 78 | 34.4 | 330 | 2 | C58931 | 1 1.45e+00 |
| 20 | 76 | 33.5 | 96 | 2 | B40957 | 1 2.84e+00 |
| 21 | 76 | 33.5 | 325 | 1 | S49358 | 1 2.84e+00 |
| 22 | 76 | 33.5 | 325 | 1 | DNW7U1 | 1 2.84e+00 |
| 23 | 76 | 33.5 | 325 | 1 | S49576 | 1 2.84e+00 |

| | | | | | | | |
|----|----|------|------|---|---------|--------------------------|----------|
| 24 | 75 | 33.5 | 33 | 1 | DNBU1 | NADH dehydrogenase (u | 1.94e-00 |
| 25 | 75 | 33.0 | 94 | 2 | A49855 | heat shock protein Gr | 3.95e-00 |
| 26 | 75 | 33.0 | 94 | 2 | A41884 | heat shock protein G | 3.95e-00 |
| 27 | 75 | 33.0 | 96 | 2 | F7152 | hypothetical protein | 3.96e-00 |
| 28 | 75 | 33.0 | 225 | 2 | F71590 | hypothetical protein | 3.96e-00 |
| 29 | 74 | 32.6 | 94 | 2 | JC1479 | heat shock protein G | 5.51e-00 |
| 30 | 74 | 32.6 | 894 | 2 | F59730 | cell wall-associated | 5.51e-00 |
| 31 | 74 | 32.6 | 1005 | 2 | JC5526 | kinase-destructive Bph | 5.51e-00 |
| 32 | 73 | 32.2 | 88 | 2 | PC1238 | heat shock protein Gr | 5.51e-00 |
| 33 | 73 | 32.2 | 94 | 2 | S12513 | chaperonin, groES, h | 5.51e-00 |
| 34 | 73 | 32.2 | 129 | 2 | T15050 | NADH dehydrogenase (u | 5.51e-00 |
| 35 | 73 | 32.2 | 404 | 2 | T05556 | pectate lyase (EC 4.2 | 7.63e-00 |
| 36 | 73 | 32.2 | 517 | 2 | T02192 | pectate lyase (EC 4.2 | 7.63e-00 |
| 37 | 72 | 31.7 | 204 | 2 | P54153 | cytochrome P450, alcohol | 1.05e-01 |
| 38 | 72 | 31.7 | 301 | 2 | H70890 | deoxycycline triphos | 1.05e-01 |
| 39 | 72 | 31.7 | 404 | 2 | S12209 | hypothetical protein | 1.05e-01 |
| 40 | 72 | 31.7 | 434 | 2 | S29612 | pectate lyase (EC 4.2 | 1.05e-01 |
| 41 | 72 | 31.7 | 438 | 2 | S43335 | pectate lyase (EC 4.2 | 1.05e-01 |
| 42 | 71 | 31.3 | 2813 | 1 | YMHU | von Willebrand factor | 1.05e-01 |
| 43 | 71 | 31.3 | 121 | 2 | 1580941 | NF-kappa B complex su | 1.45e-01 |
| 44 | 71 | 31.3 | 326 | 2 | S59131 | NADH dehydrogenase (u | 1.45e-01 |
| 45 | 71 | 31.3 | 558 | 2 | A42023 | transcription activat | 1.45e-01 |

ALIGNMENTS

```

RESULT      1
ENTRY
  TITLE      heat shock protein groES - Mycobacterium leprae
  ABBREVIATE 10K chaperonin
  ALTERNATE_NAMES
  ORGANISM   *Mycobacterium leprae*
  DATE       03-Feb-1994
  *sequence_revision 03-Feb-1994
  *text_change 26-Aug-1999
  S25180: S72997
  S25180: S72997
  S25180
  *authors   de Wit, T.F.R.; Bekelie, S.; Osland, A.; Mikk, T.L.; Hermans,
    P.W.M.; van Soelingen, D.; Drifhout, J.; Schoenlingh, R.;
    Janson, A.A.M.; Toie, J.F.R.
  *Journal   Mol. Microbiol. (1992) 6:1955-2007
  *title      Mycobactera contain two groEL genes: the second
    Mycobacterium leprae groEL gene is arranged in an operon
    with groES.
  *cross-references MIMD:92374850
  *accession  S25180
  *molecule_type DNA
  *residues    1-100
  *cross-references EMBL:Z166
  *status      The authors transferred the publication from J Mol Biol
    to this database.

```

```

RESULTS
ENTRY 4
TITLE 10K T-cell antigen - Mycobacterium leprae
ORGANISM Mycobacterium leprae
DATE 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 12-Sep-1997
ACCESSIONS JH0556 #type complete
REFERENCE JH0556
#authors Mehra, V.; Bloom, B.R.; Butardi, A.C.; Grisso, C.L.; Sneling, P.A.; Alland, D.; Convit, C.; Fan, X.; Hunter, S.W.; Brennan, P.J.; Rea, T.H.; Modlin, R.L.
#journal J. Exp. Med. (1992) 175:275-284
#title A major T cell antigen of Mycobacterium leprae is a 10-KD heat-shock cognate protein.
#cross-references NCID:92113469
#accession JH0556
#status nucleic acid sequence not shown
#molecule_type DNA
#residues 1-100 #label MEH
#cross-references EMBL:X59413
#comment This protein is a heat-shock protein.
CLASSIFICATION superfamily chaperonin groES
SUMMARY #length 100 #molecular_weight 10768 #checksum 7371
Query Match 50.2% Score 114 DB 2 Length 100
Best Local Similarity 100.0% Pred. No. 1,720-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DC 84 EFYLILSARDVLAVVSK 100
QY 1 EFYLILSARDVLAVVSK 17
RESULTS 3
ENTRY 3
TITLE Chaperonin groES - Mycobacterium tuberculosis
ALTERNATE_NAMES 10K antigen; BCG-a homolog; Cpn10; heat shock protein; 10K (hsp10); immunogenic protein BCG-a
ORGANISM Mycobacterium tuberculosis
DATE 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Jul-1995
ACCESSIONS S01381; A37166; S02727; A47292; 370747; A46481; S16040
REFERENCE S01381; A37166; S02727; A47292; 370747; A46481; S16040
#authors Baird, P.N.; Hall, L.M.C.; Coates, A.R.M.
#journal Nucleic Acids Res. (1986) 16:19347
#title A major antigen from Mycobacterium tuberculosis which is homologous to the heat shock proteins arises from a gene and the hspA gene product of Escherichia coli
#cross-references NCID:9016964
#accession S01381
#molecule_type DNA
#residues 1-100 #label BAI
#cross-references EMBL:X12598; NID:q44551; PID:q591352
REFERENCE A37166
#authors Baird, P.N.; Hall, L.M.C.; Coates, A.R.M.
#journal J. Gen. Microbiol. (1989) 135:931-939
#title Cloning and sequence analysis of the 10 kDa antigen gene of Mycobacterium tuberculosis.
#cross-references NCID:90095443
#accession A37166
#status preliminary
#molecule_type DNA
#residues 1-100 #label BA2
#cross-references G5:M25258; G3:X12593
REFERENCE S02727
#authors Shinnick, T.M.; Plikaytis, B.B.; Hyshe, A.D.; van Landingham, R.M.; Walker, L.L.
#journal Nucleic Acids Res. (1989) 17:1254
#title The Mycobacterium tuberculosis BCG-a protein has homology with the Escherichia coli GroES protein.
#cross-references NCID:89150258
#accession S02727
#status preliminary
#molecule_type DNA
#residues 1-100 #label SHI
#cross-references EMBL:X13739; NID:q44571; PID:CAA32303; PID:q591660
REFERENCE A47292
#authors Kong, T.H.; Coates, A.R.; Butcher, P.D.; Hickman, C.; Shinnick, T.M.
#journal Proc. Natl. Acad. Sci. U.S.A. (1993) 90:2639-2712
#title Mycobacterium tuberculosis expresses two chaperonin homologs.
#cross-references NCID:93219332
#accession A47292
#status preliminary
#molecule_type DNA
#residues 1-100 #label KON
#cross-references EMBL:X60350; NID:q44559; PID:CAA42908; PID:q591363
#note sequence extracted from NCBI BankOne (NCBI:128605; NCBI:P:128607)
REFERENCE A70500
#authors Cole, S.J.; Brosch, R.; Parkhill, J.; Garnier, J.; Churcher, C.; Harris, D.; Gordon, S.V.; Eigmeier, K.; Gas, S.; Barry, I.; Chillingworth, T.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Harlin, N.; Holroyd, S.; Hornsby, T.; Jagels, K.; Krogh, A.; McLean, J.; Miller, S.; Murphy, L.; Oliver, S.; Osborne, J.; Quail, M.A.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Squires, R.; Sutton, J.; Taylor, K.; Whitehead, S.; Barrett, B.G.
#journal Nature (1998) 393:537-544
#title Deciphering the biology of Mycobacterium tuberculosis from the complete genome sequence.
#cross-references NCID:98295987
#accession S70737
#status nucleic acid sequence not shown; translation not shown
#molecule_type DNA
#residues 1-100 #label GCL
#cross-references GB:277165; IS-A:123456; NID:q3261609; PID:CAA30105; PID:q55190; PID:q1449369
#experimental_source strain 1479v
REFERENCE A46481
#authors Barnes, P.F.; Mehra, V.; Kivett, P.; Fong, S.J.; Brennan, P.J.; Voca, M.S.; Minden, P.; Hopstetter, R.A.; Bloom, B.R.; Modlin, R.L.
#journal J. Immunol. (1992) 148:1845-1849
#title Immunoreactivity of a 10-kDa antigen of Mycobacterium tuberculosis.
#cross-references NCID:92171547
#accession A46481
#status preliminary
#molecule_type DNA
#residues 1-100 #label HAP
#note sequence extracted from NCBI BankOne (NCBI:P:87128)
GENETICS
#gene groES; cpn10
#start_codon GTG
#complex functional chaperonin includes 14 chains of groEL and 7 of groES
#function mediates protein folding and renaturation
#description #superfamily chaperonin groES
#keywords ATP; heat shock; molecular chaperone; stress-induced protein
#feature 2-100
#product chaperonin; status experimental; label MAT
SUMMARY #length 100 #molecular_weight 10804 #checksum 7634
Query Match 50.2% Score 114 DB 2 Length 100
Best Local Similarity 100.0% Pred. No. 1,720-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 84 EFYLILSARDVLAVVSK 100
1 EFYLILSARDVLAVVSK 17

```

```

QY 1 EBYLILSARDVLAVVSK 17

RESULT 4
ENTRY
TITLE heat shock protein cnpA - Mycobacterium leprae
ALTERNATE_NAMES heat shock protein cnpA - Mycobacterium leprae
ORGANISM Mycobacterium leprae
DATE 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change
26-Aug-1999
ACCESSIONS S72818 #type complete
REFERENCE S72550
AUTHORS Smith, D.R.; Robinson, K
SUBMISSION submitted to the FMBL Data Library, November 1993
DESCRIPTION Mycobacterium leprae cosmid B1620.
ACCESSION S72818
STATUS preliminary
#molecule_type DNA
#residues 1-169 #label SMZ
#cross-references EMBL:U000015; NID:g466931; PIDN:AA03227.1;
PID:g466935

GENETICS
#gene cnpA
#superfamily chaperonin groES
#keywords heat shock; molecular chaperone; stress-induced protein
SUMMARY #length 169 #molecular_weight 18997 #checksum 6011

Query Match 50.2% Score 114; DB 2; Length 169;
Best Local Similarity 100.0%; Pred. No. 1.72e-06;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 153 EBYLILSARDVLAVVSK 169
QY 1 EBYLILSARDVLAVVSK 17

RESULT 5
ENTRY
TITLE Chaperonin groES - Streptomyces coelicolor
ALTERNATE_NAMES Chaperonin groES - Streptomyces coelicolor
ORGANISM Streptomyces coelicolor
DATE 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change
26-Aug-1999
ACCESSIONS S37565 #type complete
REFERENCE S37564
AUTHORS Ducheene, A.M.; Kieser, K.; Hopwood, D.; Thompson, C.;
Mazodier, P.
SUBMISSION submitted to the FMBL Data Library, September 1993
DESCRIPTION Molecular characterization of two groEL genes in Streptomyces
coelicolor A3(2).
ACCESSION S37565 #preliminary
#status preliminary
#molecule_type DNA
#residues 1-132 #label DUC
#cross-references EMBL:X75206; NID:g406595; PIDN:CAA53018.1;
PID:g809756

GENETICS
#gene groES
#start_codon ATG
#superfamily chaperonin groES
#keywords heat shock; molecular chaperone; stress-induced protein
SUMMARY #length 102 #molecular_weight 10946 #checksum 4809

Query Match 46.3% Score 105; DB 2; Length 102;
Best Local Similarity 82.4%; Pred. No. 5.50e-05;
Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

DB 86 EBYLILSARDVLAVVSK 102
QY 1 EBYLILSARDVLAVVSK 17

RESULT 6
ENTRY
TITLE heat shock protein 18 - Streptomyces albus
ALTERNATE_NAMES heat shock protein groES homolog
ORGANISM Streptomyces albus
DATE 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change
12-Sep-1997
ACCESSIONS A41325
REFERENCE A41325
AUTHORS Mazodier, P.; Gugglielmi, G.; Davies, J.; Thompson, C.
JOURNAL Bacteriol. (1992) 173:7392-7396
TITLE Characterization of the groEL-like genes in Streptomyces
albus.
#cross-references MUID:92041639
#accession A41325
#molecule_type DNA
#residues 1-102 #label MAZ
#cross-references GB:M76657

GENETICS
#start_codon GTG
#superfamily chaperonin groES
CLASSIFICATION #length 102 #molecular_weight 10955 #checksum 464;
SUMMARY

Query Match 45.4% Score 103; DB 2; Length 102;
Best Local Similarity 76.5%; Pred. No. 2.43e-04;
Matches 13; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

DB 86 EBYLILSARDVLAVVSK 102
QY 1 EBYLILSARDVLAVVSK 17

RESULT 7
ENTRY
TITLE major allergen Cry j I precursor (clone pCCI-15) - Japanese
cedar
ALTERNATE_NAMES Cryptomeria japonica #common_name Japanese cedar
ORGANISM Cryptomeria japonica
DATE 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change
26-Aug-1999
ACCESSIONS JC2124 #type complete
REFERENCE JC2124
AUTHORS Sone, T.; Komatsu, N.; Shimizu, K.; Kusakabe, T.; Marikubo,
K.; Kino, K.
JOURNAL Biochem. Biophys. Res. Commun. (1994) 199:619-625
TITLE Cloning and sequencing of cDNA coding for Cry j I, a major
allergen of Japanese cedar pollen.
#cross-references MUID:94183234
#accession JC2124
#molecule_type mRNA
#residues 1-374 #label SON
#cross-references SB:025545; NID:g49443; PIDN:U00637; EMBL:U00634
#experimental_source pollen
#note the authors described carbohydrate binding site for
residue 279

CLASSIFICATION #superfamily peptidase
KEYWORDS glycoprotein; pollen
FEATURE
1-21
22-374
158-159,293,354 #domain signal sequence #status predicted #label SIG
#product major allergen Cry j I (clone pCCI-15) #status
predicted #label MAZ
#binding_site carbohydrate (Asn) (covalent) #status
predicted
SUMMARY #length 374 #molecular_weight 40702 #checksum 3592

Query Match 43.2% Score 98; DB 2; Length 374;
Best Local Similarity 86.7%; Pred. No. 1.00e-03;
Matches 13; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

DB 232 KSMKVTVAFFNOFGPN 246
QY 18 RMKVTVAFFNOFGPN 32

RESULT 8
ENTRY
TITLE JC2123 #type complete

```

```

TITLE      major allergen Cry 1 precursor (clone pCCT-2-2) - Japanese cedar
ORGANISM   #formal_name Cryptomeria japonica #common_name Japanese cedar
DATE       14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change
ACCESSIONS J02123: PC2065
REFERENCE   J02123: PC2065
#authors   Sone, T.; Komiyama, N.; Shimizu, K.; Kusakabe, T.; Morikubo, K.; Kino, K.
#journal   Biochem. Biophys. Res. Commun. (1994) 199:619-625
#title     Cloning and sequencing of cDNA coding for Cry 1, a major allergen of Japanese cedar pollen.
#cross-references MIM:941832
#accession J02123
#molecule_type mRNA
#residues  1-374 #label SN
#cross-references GB:026544; NID:q493631; PID:d1006066; PID:q493632
#experimental_source pollen
#accession PC2065
#molecule_type protein
#residues  22-53;58-81;219-232;236-258;299-307;346-372 #label S02
#note      the authors described carbohydrate binding site for residue 279
CLASSIFICATION #superfamily pectate lyase LA759
KEYWORDS      glycoprotein; pollen
FEATURE
1-21         #domain signal sequence #status predicted #label SIG\
22-374       #product major allergen Cry 1 (clone pCCT-2-2) #status predicted #label WAT\
158-191,293,354 #binding_site carbohydrate (Asn) (covalent) #status predicted
SUMMARY      #length 374 #molecular_weight 40545 #checksum 2920
Query Match 43.2% Score 98 DB 2: Length 374:
Best Local Similarity 86.7% Pred. No. 1,00e-03:
Matches 13: Conservative 1: Mismatches 1: Indels 0: Gaps 0:
DB 232 KSKVTVAFFNFGPN 246
QY 18 PRMKVTVAFFNFGPN 12

RESULT 9
ENTRY   E71621 #type complete
TITLE   ERCC-like excision repair protein PFB0160w - malaria parasite (Plasmodium falciparum)
ORGANISM #formal_name Plasmodium falciparum
DATE     13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change
ACCESSIONS E71621
REFERENCE   A71621
#authors   Gardner, M.J.; Tetlow, B.; Garavito, R.M.; Cummings, L.M.; Aravind, L.; Koonin, E.V.; Shalimov, S.; Mason, J.; Yu, K.; Fujii, C.; Pederson, J.; Shen, K.; Jing, J.; Aston, C.; Kal, Z.; Schwartz, G.C.; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, R.O.; Fraser, C.M.; Adams, M.D.; Venter, J.C.; Hoffman, S.D.; Science (1998) 282:1126-1132
#journal   Science (1998) 282:1126-1132
#title     Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
#cross-references MIM:95021743
#accession E71621
#status   preliminary; nucleic acid sequence not shown; translation not shown
#molecule_type DNA
#residues  1-242 #label GAR
#cross-references GB:A5001377; GB:A5001362; NID:q3845111; PID:q3845122; #GR:PFB0160w
#experimental_source clone 3D7
GENETICS
SUMMARY   PFB0160w
#length 242 #molecular_weight 28257 #checksum 9994

```

```

Query Match 38.8% Score 88 DB 2: Length 242:
Best Local Similarity 40.6% Pred. No. 4.27e-02:
Matches 13: Conservative 8: Mismatches 11: Indels 0: Gaps 0:
DB 45 EOYLITSLROKLPVIKKRVVYKFN:IPD 76
QY 1 EYLILSARDVLAVSKRRMKVTVAFNFGPN 32

RESULT 10
ENTRY   T00856 #type complete
TITLE   pectate lyase (EC 4.2.2.2) 120F6.14 - Arabidopsis thaliana
ORGANISM #formal_name Arabidopsis thaliana #common_name thaliana
DATE     12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change
ACCESSIONS T00856
REFERENCE   Z14206
#authors   Rounsley, S.D.; Lin, X.; Ketchum, K.A.; Crosby, M.; Brandon, R.C.; Sykes, S.M.; Kaul, S.; Mason, T.M.; Kerlavage, A.R.; Adams, M.D.; Somerville, C.R.; Venter, J.C.
#submission submitted to the C330 Data Library, March 1998
#description Arabidopsis thaliana chromosome 11 BAC 120F6 genomic sequence.
#accession T00856
#status   translated from GB/LMPL/DB02
#molecule_type DNA
#residues  1-455 #label RC0
#cross-references EMBL:AC021111; NID:q4947056; PID:AA1054511; PID:q2947
#experimental_source cultivar Columbia
GENETICS
#map-position 2
#introns 56/2: 295/3: 375/3
#exon 720F6.14
#note      #superfamily pectate lyase LA159
CLASSIFICATION #carbon-oxygen lyase
KEYWORDS
SUMMARY      #length 455 #molecular_weight 51257 #checksum 111
Query Match 37.0% Score 84 DB 2: Length 455:
Best Local Similarity 36.7% Pred. No. 1.8e-01:
Matches 11: Conservative 11: Mismatches 6: Indels 2: Gaps 2:
DB 294 DEVLPFGGTHK-DVIDKK-MQITVAFNFG 321
QY 1 EYLILSARDVLAVSKRRMKVTVAFNFG 30

RESULT 11
ENTRY   E00605 #type complete
TITLE   chaperonin 60kDa - Synecchococcus sp. (strain 6805)
ALTERNATE_NAMES Heat shock protein, hsp 60
ORGANISM #formal_name Synecchococcus sp.
DATE     30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change
ACCESSIONS S10836; S09619
REFERENCE   S07286
#authors   Cozens, A.L.; Walker, J.E.
#journal   J. Mol. Biol. (1987) 194:359-383
#title     The organization and sequence of the genes for ATP synthase subunits in the cyanobacterium Synecchococcus 6801. Support for an endosymbiotic origin of chloroplasts.
#cross-references MIM:87311713
#accession S10836
#molecule_type DNA
#residues  1-103 #label CO2
#cross-references EMBL:X05925; NID:q48021; PID:CAA29361.1; PID:q48023
#experimental_source functional chaperonin includes 14 chains of groEL and 7 of groES
FUNCTION
#description mediates protein folding and renaturation.
CLASSIFICATION #superfamily chaperonin groES

```

KEYWORDS ATP: heat shock: molecular chaperone; stress-induced protein
 FEATURE 2-97 *product chaperonin groES *status predicted *label MAT
 SUMMARY *length: 103 *molecular-weight 10811 *checksum 6496

Query Match 35.2% Score 80; DB 1; Length 103;
 Best Local Similarity 50.0%; Pred. No. 7.31e-01;
 Matches 8; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

DB 88 DQYVLLSEKDLAVVA 103
 QY 1 EYVLLSARDVLAVS 16

RESULT 12 *type complete
 ENTRY groES protein - Synecococcus sp. (strain PCC 7942)
 TITLE *formal_name Synecococcus sp.
 ORGANISM 12-Apr-1991; *sequence_revision 12-Apr-1991 *text_change
 DATE 26-Aug-1995

ACCESSIONS A36721
 REFERENCE Webb, R.; Reddy, K.J.; Sherman, L.A.
 J. Bacteriol. (1990) 172:5079-5088
 *journal Regulation and sequence of the Synecococcus sp. strain PCC
 7942 drbSL operon, encoding a cyanobacterial chaperonin.
 *cross-references NCBI:90368561

*accession A36721
 *status preliminary
 *molecule_type DNA
 *residues 1-103 *label WEB
 *cross-references GB:M58751; NID:q54519; PIDN:AA27313.1; PID:q154520
 CLASSIFICATION *superfamily chaperonin groES
 SUMMARY *length 103 *molecular-weight 10742 *checksum 6552

Query Match 35.2% Score 80; DB 2; Length 103;
 Best Local Similarity 50.0%; Pred. No. 7.31e-01;
 Matches 8; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

DB 88 DQYVLLSEKDLAVVA 103
 QY 1 EYVLLSARDVLAVS 16

RESULT 13 *type complete
 ENTRY NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 1 -
 TITLE Pedinorcas minor mitochondrion (SGC3)
 ORGANISM *formal_name mitochondrion: Pedinomonas minor
 DATE 16-Jul-1999; *sequence_revision 16-Jul-1999 *text_change
 13-Aug-1995

ACCESSIONS T11320
 REFERENCE Turmel, M.; Lemieux, C.; Burger, G.; Lang, B.F.; Odis, C.;
 Plante, I.; Gray, M.W.
 submitted to the EMBL Data Library, December 1998
 *description The complete mitochondrial DNA sequences of Nephroselmis
 olivacea and Pedinomonas minor: two radically different
 evolutionary patterns within the green algae.

*accession T11320
 *status preliminary; translated from GB/EMBL/DBJ
 *molecule_type DNA
 *residues 1-310 *label TJR
 *cross-references EMBL:AF116775; NID:q4378756; PID:q4378774;
 PIDN:AD19672.1

GENETICS
 *gene nadl
 *genome mitochondrion;
 *genetic_code SGC3

CLASSIFICATION *superfamily NADH dehydrogenase (ubiquinone) chain 1
 KEYWORDS mitochondrion; NAD; oxidoreductase
 SUMMARY *length 310 *molecular-weight 34938 *checksum 2458

Query Match 35.2% Score 80; DB 2; Length 310;
 Best Local Similarity 41.7%; Pred. No. 7.31e-01;
 Matches 10; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

DB 28 FLVLIERKVIAYIQRGGPVFGP 310
 QY 3 YLILSARDVLAVVSKRRMKVIVAF 26

RESULT 14 *type complete
 ENTRY NADH dehydrogenase (ubiquinone) (EC 1.6.5.3) chain 1 -
 TITLE Prototheca wickerhamii mitochondrion
 ORGANISM *formal_name mitoch. orgion: Prototheca wickerhamii
 DATE 16-Jul-1999; *sequence_revision 16-Jul-1999 *text_change
 13-Aug-1995

ACCESSIONS T11938
 REFERENCE Wolff, G.; Plante, I.; Lang, B.F.; Kueck, T.; Burger, G.
 J. Mol. Biol. (1997) 237:75-86
 *journal Complete sequence of the mitochondrial DNA of the chlorophyte
 alga Prototheca wickerhamii. Gene content and genome
 organization.
 *cross-references NCBI:94180393

*accession T11938
 *status preliminary; translated from GB/EMBL/DBJ
 *molecule_type DNA
 *residues 1-328 *label WGL
 *cross-references EMBL:U02970; NID:q467843; PID:q467870;
 PIDN:AA012657.1

*experimental_source strain 263-11
 GENETICS
 *genome mitochondrion
 *note nadl

CLASSIFICATION *superfamily NADH dehydrogenase (ubiquinone) chain 1
 KEYWORDS mitochondrion; NAD; oxidoreductase
 SUMMARY *length 328 *molecular-weight 36437 *checksum 9332

Query Match 35.2% Score 80; DB 2; Length 328;
 Best Local Similarity 33.3%; Pred. No. 7.31e-01;
 Matches 10; Conservative 8; Mismatches 11; Indels 1; Gaps 1;

DB 25 FLVLAERKVLAKMCRKGNVVGSLIFGP 54
 QY 3 YLILSARDVLAVVSKRRMKVIVAFVQFGP 31

RESULT 15
 ENTRY BMVW7B *type complete
 TITLE chaperonin groES, Mycobacterium bovis
 ALTERNATE_NAMES heat shock protein, hsp (hsp10); immunogenic protein MP857
 ORGANISM *formal_name Mycobacterium bovis
 DATE 31-Mar-1991; *sequence_revision 31-Mar-1991 *text_change
 16-Jul-1999

ACCESSIONS SC1784; A37534; C60278
 REFERENCE SO1784
 *authors Yanaguchi, R.; Matsuo, K.; Yamazaki, A.; Nishi, S.; Teraseka,
 K.; Yamada, T.
 *journal FEBS Lett. (1988) 240:115-117
 *title Immunogenic protein MP857 from Mycobacterium bovis H37:
 molecular cloning, nucleotide sequence and expression.

*cross-references NCBI:89052866
 *accession SO1784
 *molecule_type DNA
 *residues 1-100 *label YAM
 *cross-references EMBL:X11970; NID:q44190; PIDN:CAA32149.1; PID:q581313
 *accession A37534

*molecule_type protein
 *residues 2-21 *label AM2
 *accession A60278
 *authors Flis, T.; Costopoulos, C.; Radford, A.J.; Bacic, A.; Wood,
 P.R.

*journal Infect. Immun. (1991) 59:800-807

#title Purification and Characterization of major antigens from a
Mycobacterium bovis culture filtrate.

#cross-references MUID:9:147217

#accession C60278

#molecule-type protein

#residues 2:15, E, 17-20 #label F1F

GENETICS

#gene

#start_codon GTG

CLASSIFICATION #superfamily chaperonin groES

KEYWORDS heat shock; molecular: chaperone; stress-induced protein

FEATURE

2-100

SUMMARY #product chaperonin groES #status predicted #label MAI

#length 100 #molecular_weight 10950 #checksum 7378

Query Match 34.8% Score 79: DB 1: length 100:

Best Local Similarity 84.6% Pred. No. 1: C: +20:

Matches 11: Conservative 2: Mismatches 0: Indels 0: Gaps 0:

EE 84 EYVILSARDVWG 96

1 1 1 1 1 1 1 1 1 1

QY 1 EYVILSARDVLA 13

Search completed: Mon Jun 19 16:12:12 2000

for time 1.9 secs.

[REDACTED]

(TM)

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (C) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsearch_pp protein - protein database search using Smith-Waterman algorithm
Run on: Mon Jun 19 16:10:49 2000: MaxPar time 3.59 Seconds
Tabular output not generated. 277.723 Million cell updates/sec

Title: >US-09-142-524A-10
Description: (1-32) from US09142524A.pep
Perfect Score: 227
Sequence: 1 EEXYLLSARDLVAVSVKRRKKV-VAFNQFGPN 32

Scoring table: PWM 15C
Gap 11

Searched: 83857 seqs. 30454973 residues
Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: swiss-prot38
1:swissprot

Statistics: Mean 35.264: Variance 53.818: scale 0.655

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

| Result No. | Score | Query Match | Locatn | DP | ID | Description | Prod. No. |
|------------|-------|-------------|--------|----|------------|-------------------------|-----------|
| 1 | 114 | 50.2 | 49 | 1 | CH10_MYCTU | 10 KD CHAPERONIN (PROT | 7.57e-08 |
| 2 | 114 | 50.2 | 49 | 1 | CH10_MYCTU | 10 KD CHAPERONIN (PROT | 7.57e-08 |
| 3 | 114 | 50.2 | 99 | 1 | CH10_MYCAV | 10 KD CHAPERONIN (PROT | 7.57e-08 |
| 4 | 105 | 45.3 | 102 | 1 | CH10_STRCO | 10 KD CHAPERONIN (PROT | 4.32e-06 |
| 5 | 103 | 45.4 | 102 | 1 | CH10_STRAL | 10 KD CHAPERONIN (PROT | 1.06e-05 |
| 6 | 98 | 43.2 | 374 | 1 | SRP_CRYJA | SUG1: BASIC PROTEIN PRE | 9.54e-05 |
| 7 | 80 | 35.2 | 100 | 1 | CH10_IETH | 10 KD CHAPERONIN (PROT | 1.60e-01 |
| 8 | 80 | 35.2 | 103 | 1 | CH10_SYNP6 | 10 KD CHAPERONIN (PROT | 1.60e-01 |
| 9 | 80 | 35.2 | 103 | 1 | CH10_SYNP7 | 10 KD CHAPERONIN (PROT | 1.60e-01 |
| 10 | 79 | 34.8 | 99 | 1 | CH10_MYCBO | 10 KD CHAPERONIN (PROT | 2.35e-01 |
| 11 | 79 | 34.8 | 869 | 1 | PLC1_YEAST | 1-PHOSPHATIDYLINOSITOL | 2.35e-01 |
| 12 | 78 | 34.4 | 102 | 1 | CH10_SYNVU | 10 KD CHAPERONIN (PROT | 3.44e-01 |
| 13 | 78 | 34.4 | 103 | 1 | CH10_CVAPA | 10 KD CHAPERONIN (PROT | 3.44e-01 |
| 14 | 78 | 34.4 | 230 | 1 | NTPO_ENTHR | V-TYPE SODIUM ATP SYNT | 3.44e-01 |
| 15 | 76 | 33.5 | 96 | 1 | CH10_LEPIN | 10 KD CHAPERONIN (PROT | 7.32e-01 |
| 16 | 76 | 33.5 | 325 | 1 | NULM_WHEAT | NADH-UBIQUINONE OXIDOR | 7.32e-01 |
| 17 | 75 | 33.5 | 325 | 1 | NULM_PETHY | NADH-UBIQUINONE OXIDOR | 7.32e-01 |
| 18 | 75 | 33.5 | 331 | 1 | NULM_CENBE | NADH-UBIQUINONE OXIDOR | 7.32e-01 |
| 19 | 75 | 33.5 | 94 | 1 | CH10_BACST | 10 KD CHAPERONIN (PROT | 1.05e+00 |
| 20 | 75 | 33.5 | 94 | 1 | CH10_BACSU | 10 KD CHAPERONIN (PROT | 1.05e+00 |
| 21 | 74 | 32.6 | 94 | 1 | CH10_BACP3 | 10 KD CHAPERONIN (PROT | 1.54e-00 |
| 22 | 74 | 32.6 | 104 | 1 | CH11_BRALA | 10 KD CHAPERONIN (PR | 1.54e-00 |
| 23 | 74 | 32.6 | 387 | 1 | VFD_SPP2 | GENE D PROTEIN (GPD) | 1.54e-00 |

| RESULT ID | CH10_MYCTU | STANDARD | PRTE | 99 AA |
|-----------|--|----------|------|-------|
| AC | FC9521 | | | |
| DI | 01-MAR-1989 (Rel. 10, Created) | | | |
| DI | 15-FEB-2000 (Rel. 39, Last sequence update) | | | |
| DI | 15-FEB-2000 (Rel. 39, Last annotation update) | | | |
| DE | 10 KD CHAPERONIN (PROTEIN GPN:0) (PROTEIN GPN:0) (BAG-A HEALISER-1X | | | |
| DE | PROTEIN (10 KD ANTIGEN) | | | |
| GN | MCPB OR GPN:0 OR CPN10 OR RV34:18C OR MTCY78.11. | | | |
| CS | Mycobacterium tuberculosis | | | |
| OC | Bacteria: Firmicutes: Actinobacteria: Actinobacteridae: | | | |
| OC | Actinomycetales: Corynebacteriaceae: Mycobacteriaceae: Mycobacterium | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN-H37RV: | | | |
| RX | MEDLINE: 89016584 | | | |
| RA | Baird P.N., Hall L.M., Coates A.R.M.: | | | |
| RA | "A major antigen from Mycobacterium tuberculosis which is homologous | | | |
| R: | to the heat shock proteins groES from E. coli and the hspA gene | | | |
| R: | product of Coxidia burnetii" | | | |
| PL | Nucleic Acids Res. 16:9047-9047(1989) | | | |
| PN | [2] | | | |
| PP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN-H37RV: | | | |
| RX | MEDLINE: 90095443 | | | |
| RA | Baird P.N., Hall L.M., Coates A.R.M.: | | | |
| R: | "Cloning and sequence analysis of the 10 kDa antigen gene of | | | |
| R: | Mycobacterium tuberculosis" | | | |
| PL | J. Gen. Microbiol. 135:931-935(1989) | | | |
| PN | [3] | | | |
| PP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN-ERDMANN: | | | |
| RX | MEDLINE: 89160258 | | | |
| RA | Shinnick T.M., Palkayis B.P., Hyshe A.D., van Landingham R.M.: | | | |
| RA | Walker L.L.: | | | |
| RT | "The M. tuberculosis protein has homology with the | | | |
| RT | Escherichia coli GroES protein" | | | |
| RL | Nucleic Acids Res. 17:1254-1254(1989) | | | |
| RN | [4] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN-ERDMANN: | | | |
| RX | MEDLINE: 93219332 | | | |
| RA | Kong T.H., Coates A.R.M., Butcher P.D., Hickman C.J., Shinnick T.M.: | | | |
| RA | "Mycobacterium tuberculosis expresses two chaperonins" | | | |
| PL | Proc. Natl. Acad. Sci. U.S.A. 90:2608-2612(1993) | | | |
| RN | [5] | | | |
| RP | SEQUENCE FROM N.A. | | | |

PC STRAIN-H37RV;
 RX MEDLINE: 94295987;
 RA Cole S.L., Brosch K., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigemeier K., Gas S., Barry C.F. III, Tekala F.,
 RA Radcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Deakin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jaisels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne C., Quail M.A., Rajandream K.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares K., Sulston J.E.,
 RA Taylor K., Whitehead S., Barrall B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RA complete genome sequence.";
 PL Nature 393:537-544(1998).
 RN [6];
 RP SEQUENCE OF 1-15;
 RX MEDLINE: 92176646;
 RA Barnes P.F., Mehra V., Riviere B., Ford S.J., Brennan P.J.,
 RA Voegtlin M.S., Minden P., Houghton R.A., Bloom B.R., Modlin R.L.;
 RA "Immunoreactivity of a 10-kDa antigen of Mycobacterium tuberculosis.";
 RL J. Immunol. 145:1835-1840(1992).
 CC -- FUNCTION: BINDS TO CPN60 IN THE PRESENCE OF MG-ATP AND SUPPRESSES
 CC THE ATPASE ACTIVITY OF THE LATTER
 CC -- SUBUNIT: HEPTAMER OF 7 SUBUNITS ARRANGED IN A KING
 CC (BY SIMILARITY).
 CC -- SIMILARITY: BELONGS TO THE GROES CHAPERONIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sb.ch.)
 DR EMBL: X60559; CAA42908.1;
 DR EMBL: X12598; CAA31111.1;
 DR EMBL: M25258; AAA23343.1;
 DR EMBL: X13739; AAA23333.1;
 DR EMBL: Z77165; JAB01605.1;
 DR PIR: S02727; BVNBYA;
 DR PIR: A37166; A37166;
 DR PIR: A47292; A47292;
 LR PRINTS: PR00297; CHAPERONIN10.
 LR PROSITE: PS00668; CHAPERONINS_CPN10: 1.
 LR PRAM: PF00186; cpn10: 1;
 LR THERMOLIST: hv3418c: 1;
 KW Chaperone: Antigen: Heat shock.
 FT INIT_MET 0
 FT SEQUENCE 99 AA: 10669 MW: 106288 DGPVAFS CRK64;
 Query Match 50.2% Score 114; DB 1; Length 99;
 Best Local Similarity 100.0%; Pred. No. 7.07e-08;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 83 EYILSARDVLAVSK 99
 QY 1 EYILSARDVLAVSK 17
 RESULT 2
 ID CH10_MYCLE STANDARD; PRI: 99 AA.
 AC P24301;
 DT 01-MAR-1992 (Rel. 21; Created)
 DT 01-MAR-1992 (Rel. 21; Last sequence update)
 DT 15-DEC-1998 (Rel. 37; Last annotation update)
 DE 10 KD CHAPERONIN (PROTEIN CPN10) (PROTEIN GROES) (10 KD ANTIGEN).
 GN XOPB OR GROES OR CHPA OR B1620_C1_227 OR B229_C3_247.
 OS Mycobacterium leprae.
 OC Bacteria: Firmicutes; Actinobacteria: Actinobacteridae;
 OC Actinomycetales: Corynebacteriaceae: Mycobacteriaceae: Mycobacterium.
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-24.
 RX MEDLINE: 9-113469.

RA Mehra V.L., Bloom B.R., Bajardi A.C., Grisso A.C., Steiling P.A.,
 RA Alland D., Convit C., Fan X., Hunter S.W., Brennan P., Rea I.E.
 RA Modlin R.L.;
 RA "A major T cell antigen of Mycobacterium leprae is a 10-kD heat-shock
 RA cognate protein.";
 RL J. Exp. Med. 175:275-284(1992).
 RN [2];
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 92374850;
 RA de Wit T.F.R., Bekelie S., Osland A., Miko T.L., Hermans P.W.M.,
 RA van Soelingen D., Drijfhout J., Schrevelink R., Jansen A.A.M.,
 RA Thole J.E.R.;
 RA "Mycobacteria contain two groEL genes, the second Mycobacterium
 RA leprae groEL gene is arranged in an operon with groES.";
 RL Mol. Microbiol. 6:1995-2007(1992).
 RN [3];
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 92374850;
 RA Smith D.R., Robison K.;
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBS databases.
 RN [4];
 RP X-RAY CRYSTALLOGRAPHY (3.5 ANGSTROMS).
 RX MEDLINE: 96138402;
 RA Wende S.C., Mehra V., Bloom B.R., Hol W.G.J.;
 RA "Structure of the heat shock protein chaperonin-10 of Mycobacterium
 RA leprae.";
 RL Science 271:203-207(1996).
 RN [5];
 RP EXTRATUM.
 RA Wende S.C., Mehra V., Bloom B.R., Hol W.G.J.;
 RL Science 271:1655-1655(1996).
 CC -- FUNCTION: BINDS TO CPN60 IN THE PRESENCE OF MG-ATP AND SUPPRESSES
 CC THE ATPASE ACTIVITY OF THE LATTER.
 CC -- SUBUNIT: HEPTAMER OF 7 SUBUNITS ARRANGED IN A KING
 CC (BY SIMILARITY: BELONGS TO THE GROES CHAPERONIN FAMILY).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sb.ch.)
 DR EMBL: X59413; NOT_ANNOTATED_CDS;
 DR EMBL: Z11665; NOT_ANNOTATED_CDS;
 DR EMBL: J00015; AAC4327.1;
 DR EMBL: J00020; AAA1731.1;
 DR PIR: JHC556; JHC556;
 DR PIR: S25189; S25189;
 DR POS: LEPR_11-AN-1;
 DR PRINTS: PR00297; CHAPERONIN10.
 DR PROSITE: PS00668; CHAPERONINS_CPN10: 1;
 DR PRAM: PF00186; cpn10: 1;
 KW Chaperone: Antigen: Heat shock; 3D-structure.
 FT INIT_MET 0
 FT SEQUENCE 99 AA: 10669 MW: 6028889633F7DEB CRK64;
 Query Match 50.2% Score 114; DB 1; Length 99;
 Best Local Similarity 100.0%; Pred. No. 7.07e-08;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 DB 83 EYILSARDVLAVSK 99
 QY 1 EYILSARDVLAVSK 17
 RESULT 3
 ID CH10_MYCAV STANDARD; PRI: 99 AA.
 AC O86017;
 DT 15-FEB-2000 (Rel. 39; Created)
 DT 15-FEB-2000 (Rel. 39; Last sequence update)
 DT 15-FEB-2000 (Rel. 39; Last annotation update)
 DE 10 KD CHAPERONIN (PROTEIN CPN10) (PROTEIN GROES) (10 KD ANTIGEN).
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-24.

GN MOPB OR GROES
 CS Mycobacterium avium, and Mycobacterium paratuberculosis.
 CC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 CC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=M.AVIUM; STRAIN=485 TYPE 21;
 RA Creti R., Pieterbong R., Fattorini L., Orefici G.;
 RT "Sequence of the groEL operon of Mycobacterium avium comprising the
 RT gene encoding the cpn10 protein and a portion of the gene encoding the
 RI cpn60-1 protein";
 R1 Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=M.PARATUBERCULOSIS, M.AVIUM; STRAIN=ATCC 19698, ATCC 25291;
 RA Cobb A.J., Fitzhugham R.;
 RI "The GROES antigens of Mycobacterium avium and Mycobacterium
 RI paratuberculosis";
 R1 Vet. Microbiol. 0:0-0(1999);
 RL VET. MICROBIOL. 0:0-0(1999);
 CC -!- FUNCTION: BINDS TO CPN60 IN THE PRESENCE OF MG-ATP AND SUPPRESS
 CC THE ATPASE ACTIVITY OF THE LATTER (BY SIMILARITY).
 CC -!- SUBUNIT: HEPTAMER OF 7 SUBUNITS ARRANGED IN A RING (BY
 CC SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE GROES CHAPERONIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: AF079544; AAC1921.1; 1;
 DR EMBL: AF071829; AAD23277.1; 1;
 DR EMBL: AF071828; AAD23276.1; 1;
 DR PFAM: PF00166; CPN10; 1;
 DR PRINTS: PR00327; CHAPERONIN10.
 DR PROSITE: PS00681; CHAPERONINS_CPN10; 1.
 KW Chaperone; Antigen; Heat shock.
 FT INIT_MET 0 BY SIMILARITY.
 ST SEQUENCE 99 AA: 10617 MW: 6804457405310C2 CRC64:

 Query Match 50.2% Score 114; DB 1; Length 99;
 Best Local Similarity 100.0%; Pred No. 7, 27e-08;
 Matches 17; Conservative 3; Mismatches 0; Indels 0; Gaps 0;
 DB 83 EBYLVLSARDVLAIVSK 99
 QY 1 EBYLVLSARDVLAIVSK 17

 RESULT 4
 ID: CH10_STRCO STANDARD: PRI: 102 AA.
 AC P40672;
 CC 01-FEB-1995 (Rel. 31, Created);
 DT 01-FEB-1995 (Rel. 31, Last sequence update);
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 CE 10 K3 CHAPERONIN (PROTEIN CPN10) (PROTEIN GROES).
 GN GROES OR SC664.39
 CC Streptomyces coelicolor, and Streptomyces lividans.
 CC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 CC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.COELICOLOR; STRAIN=A3(2) / J1501;
 RX MEDLINE: 94299177.
 RA Suzanne A.M., Kleser H.M., Hopwood D.A., Thompson C.J., Mazodier P.;
 RT "Characterization of two groEL genes in Streptomyces coelicolor
 R1 A3(2)";
 R1 Gene 144:97-101(1994).
 RN [2]
 RP SEQUENCE FROM N.A.

RC SPECIES=S.COELICOLOR; STRAIN=A3(2);
 RA Saunders D., Harris D., Parkhill J., Barrell B.G., Rajandream M.A.,
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC SPECIES=S.LIVIDANS;
 RX MEDLINE: 98048481.
 RA de Leon P., Marco S., Iseegas C., Marina A., Carrascosa J.L.,
 RA Mellado R.P.;
 RT "Streptomyces lividans groES, groEL1 and groEL2 genes";
 R1 Microbiology 143:3583-3571(1997);
 RL MICROBIOLOGY 143:3583-3571(1997);
 CC -!- FUNCTION: BINDS TO CPN60 IN THE PRESENCE OF MG-ATP AND SUPPRESSES
 CC THE ATPASE ACTIVITY OF THE LATTER.
 CC -!- SUBUNIT: HEPTAMER OF 7 SUBUNITS ARRANGED IN A RING
 CC (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE GROES CHAPERONIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: X75206; CAA53018.1; 1;
 DR EMBL: AL031317; CAA20417.1; 1;
 DR EMBL: X95970; CAA65224.1; 1;
 DR PIR: S37565; S37565.
 DR PRINTS: PR00297; CHAPERONIN.
 DR PROSITE: PS00681; CHAPERONINS_CPN10; 1.
 DR PFAM: PF00166; CPN10; 1.
 KW Chaperone.
 ST SEQUENCE 102 AA: 10946 MW: 6159FA68014AC87C CRC64:

 Query Match 46.3% Score 105; DB 1; Length 102;
 Best Local Similarity 82.4%; Pred No. 4, 32e-06;
 Matches 14; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 DB 86 EBYLVLSARDVLAIVSK 102
 QY 1 EBYLVLSARDVLAIVSK 17

 RESULT 5
 ID: CH10_STRAL STANDARD: PRI: 102 AA.
 AC Q00769;
 CC 01-APR-1993 (Rel. 25, Created);
 DT 01-APR-1993 (Rel. 25, Last sequence update);
 DT 01-FEB-1996 (Rel. 33, Last annotation update)
 CE 10 K3 CHAPERONIN (PROTEIN CPN10) (PROTEIN GROES).
 GN GROES.
 CC Streptomyces albus G.
 CC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 CC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 92041639.
 RA Mazodier P., Guglielmi G., Davies J., Thompson C.J.;
 RT "Characterization of the groEL-like genes in Streptomyces albus G";
 R1 J. Bacteriol. 173:7382-7386(1992);
 CC -!- FUNCTION: BINDS TO CPN60 IN THE PRESENCE OF MG-ATP AND SUPPRESSES
 CC THE ATPASE ACTIVITY OF THE LATTER.
 CC -!- SUBUNIT: HEPTAMER OF 7 SUBUNITS ARRANGED IN A RING
 CC (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO THE GROES CHAPERONIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)

DR EMBL: D45880; RAA08298.1; 1;
 DR EMBL: U29483; AAA83440.1; 1;
 DR PIR: B39313; B39313;
 DR PRINTS: PR0297; CHAPERONIN10.
 DR PROSITE: PS00681; CHAPERONINS_CPN10: 1.
 DR PFAM: PF00166; cpn10; 1.
 KW Chaperone.
 FT INIT_MET 0 0
 FT 19 19 R -> K (IN REF. 3).
 FT 83 86 DGE -> APRRT (IN REF. 2).
 FT 83 86
 SQ SEQUENCE 100 AA; 10865 MW; 452F3DD9CCCB0FA CRC64;

Query Match 35.2% Score 80; DB 1; Length 100;
 Best Local Similarity 73.3%; Pred. No. 1.60e-01;
 Matches 11; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

DB 85 EYVLLSERDLAVL 99
 : EYVLLSARDVLAV 15
 : EYVLLSARDVLAV 15

RESULT 8
 ID CH10_SYN6 STANDARD: PRT: 103 AA.
 AC P07889;
 DT 01-AUG-1989 (Rel. 08, Created)
 DI 01-AUG-1989 (Rel. 08, Last sequence update)
 DE 01-FEB-1996 (Rel. 33, Last annotation update)
 DE 10 KD CHAPERONIN (PROTEIN CPN10) (PROTEIN GROES).
 GN GROES.

OS Synchococcus sp. (strain PCC 6301).
 OC Bacteria; Cyanobacteria; Chroococcales; Synchococcus.

PN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 8731173

RA Coziers A.L., Walker J.E.;
 RI "The organization and sequence of the genes for ATP synthase subunits
 RI in the cyanobacterium Synchococcus 6301. Support for an
 RI endosymbiotic origin of chloroplasts."

RC J. Mol. Biol. 194:359-383(1987).
 RN [2]

RP SIMILARITY TO CHAPERONINS.
 RX MEDLINE: 89366572.

RA Cookson M.C., Baird P.N., Hall L.M., Coates A.R.M.;
 RI "Identification of two unknown reading frames in Synchococcus 6301
 RI as homologues of the 11k and 65k antigen genes of Mycobacterium
 RI tuberculosis and related heat shock genes in E. coli and Coxsiella
 RI burnetii."

RC Nucleic Acids Res. 17:6392-6392(1989).
 RN [1]

RP THE ATPASE ACTIVITY OF THE LATER.
 RX MEDLINE: 89366572.

RA Cookson M.C., Baird P.N., Hall L.M., Coates A.R.M.;
 RI "Identification of two unknown reading frames in Synchococcus 6301
 RI as homologues of the 11k and 65k antigen genes of Mycobacterium
 RI tuberculosis and related heat shock genes in E. coli and Coxsiella
 RI burnetii."

RC Nucleic Acids Res. 17:6392-6392(1989).
 RN [1]

RP THE ATPASE ACTIVITY OF THE LATER.
 RX MEDLINE: 89366572.

RA Cookson M.C., Baird P.N., Hall L.M., Coates A.R.M.;
 RI "Identification of two unknown reading frames in Synchococcus 6301
 RI as homologues of the 11k and 65k antigen genes of Mycobacterium
 RI tuberculosis and related heat shock genes in E. coli and Coxsiella
 RI burnetii."

RC Nucleic Acids Res. 17:6392-6392(1989).
 RN [1]

RP THE ATPASE ACTIVITY OF THE LATER.
 RX MEDLINE: 89366572.

RA Cookson M.C., Baird P.N., Hall L.M., Coates A.R.M.;
 RI "Identification of two unknown reading frames in Synchococcus 6301
 RI as homologues of the 11k and 65k antigen genes of Mycobacterium
 RI tuberculosis and related heat shock genes in E. coli and Coxsiella
 RI burnetii."

RC Nucleic Acids Res. 17:6392-6392(1989).
 RN [1]

RP THE ATPASE ACTIVITY OF THE LATER.
 RX MEDLINE: 89366572.

RA Cookson M.C., Baird P.N., Hall L.M., Coates A.R.M.;
 RI "Identification of two unknown reading frames in Synchococcus 6301
 RI as homologues of the 11k and 65k antigen genes of Mycobacterium
 RI tuberculosis and related heat shock genes in E. coli and Coxsiella
 RI burnetii."

RC Nucleic Acids Res. 17:6392-6392(1989).
 RN [1]

RP THE ATPASE ACTIVITY OF THE LATER.
 RX MEDLINE: 89366572.

RA Cookson M.C., Baird P.N., Hall L.M., Coates A.R.M.;
 RI "Identification of two unknown reading frames in Synchococcus 6301
 RI as homologues of the 11k and 65k antigen genes of Mycobacterium
 RI tuberculosis and related heat shock genes in E. coli and Coxsiella
 RI burnetii."

RC Nucleic Acids Res. 17:6392-6392(1989).
 RN [1]

RP THE ATPASE ACTIVITY OF THE LATER.
 RX MEDLINE: 89366572.

RA Cookson M.C., Baird P.N., Hall L.M., Coates A.R.M.;
 RI "Identification of two unknown reading frames in Synchococcus 6301
 RI as homologues of the 11k and 65k antigen genes of Mycobacterium
 RI tuberculosis and related heat shock genes in E. coli and Coxsiella
 RI burnetii."

RC Nucleic Acids Res. 17:6392-6392(1989).
 RN [1]

Query Match 35.2% Score 80; DB 1; Length 103;
 Best Local Similarity 50.3%; Pred. No. 1.60e-01;
 Matches 8; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

DB 88 DDYVLLSEKDLAVVA 103
 : EYVLLSARDVLAV 16
 : EYVLLSARDVLAV 16

RESULT 9
 ID CH10_SYN7 STANDARD: PRT: 103 AA.
 AC P22880;
 DT 01-AUG-1991 (Rel. 19, Created)
 DI 01-AUG-1991 (Rel. 19, Last sequence update)
 DE 01-FEB-1996 (Rel. 33, Last annotation update)
 DE 10 KD CHAPERONIN (PROTEIN CPN10) (PROTEIN GROES).
 GN GROES.

OS Synchococcus sp. (strain PCC 7942).
 OC Bacteria; Cyanobacteria; Chroococcales; Synchococcus.

PN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 90368561.

RA Webb R., Reddy K.J., Sherman L.A.;
 RI "Regulation and sequence of the Synchococcus sp. strain PCC 7942
 RI groESL operon, encoding a cyanobacterial chaperonin."

RC J. Bacteriol. 172:5079-5080(1990).
 RN [1]

RP THE ATPASE ACTIVITY OF THE LATER.
 RX MEDLINE: 90368561.

RA Webb R., Reddy K.J., Sherman L.A.;
 RI "Regulation and sequence of the Synchococcus sp. strain PCC 7942
 RI groESL operon, encoding a cyanobacterial chaperonin."

RC J. Bacteriol. 172:5079-5080(1990).
 RN [1]

RP THE ATPASE ACTIVITY OF THE LATER.
 RX MEDLINE: 90368561.

RA Webb R., Reddy K.J., Sherman L.A.;
 RI "Regulation and sequence of the Synchococcus sp. strain PCC 7942
 RI groESL operon, encoding a cyanobacterial chaperonin."

RC J. Bacteriol. 172:5079-5080(1990).
 RN [1]

RP THE ATPASE ACTIVITY OF THE LATER.
 RX MEDLINE: 90368561.

RA Webb R., Reddy K.J., Sherman L.A.;
 RI "Regulation and sequence of the Synchococcus sp. strain PCC 7942
 RI groESL operon, encoding a cyanobacterial chaperonin."

RC J. Bacteriol. 172:5079-5080(1990).
 RN [1]

RP THE ATPASE ACTIVITY OF THE LATER.
 RX MEDLINE: 90368561.

RA Webb R., Reddy K.J., Sherman L.A.;
 RI "Regulation and sequence of the Synchococcus sp. strain PCC 7942
 RI groESL operon, encoding a cyanobacterial chaperonin."

RC J. Bacteriol. 172:5079-5080(1990).
 RN [1]

RP THE ATPASE ACTIVITY OF THE LATER.
 RX MEDLINE: 90368561.

RA Webb R., Reddy K.J., Sherman L.A.;
 RI "Regulation and sequence of the Synchococcus sp. strain PCC 7942
 RI groESL operon, encoding a cyanobacterial chaperonin."

RC J. Bacteriol. 172:5079-5080(1990).
 RN [1]

RP THE ATPASE ACTIVITY OF THE LATER.
 RX MEDLINE: 90368561.

RA Webb R., Reddy K.J., Sherman L.A.;
 RI "Regulation and sequence of the Synchococcus sp. strain PCC 7942
 RI groESL operon, encoding a cyanobacterial chaperonin."

RC J. Bacteriol. 172:5079-5080(1990).
 RN [1]

RP THE ATPASE ACTIVITY OF THE LATER.
 RX MEDLINE: 90368561.

RA Webb R., Reddy K.J., Sherman L.A.;
 RI "Regulation and sequence of the Synchococcus sp. strain PCC 7942
 RI groESL operon, encoding a cyanobacterial chaperonin."

RC J. Bacteriol. 172:5079-5080(1990).
 RN [1]

RP THE ATPASE ACTIVITY OF THE LATER.
 RX MEDLINE: 90368561.

RA Webb R., Reddy K.J., Sherman L.A.;
 RI "Regulation and sequence of the Synchococcus sp. strain PCC 7942
 RI groESL operon, encoding a cyanobacterial chaperonin."

RC J. Bacteriol. 172:5079-5080(1990).
 RN [1]

RP THE ATPASE ACTIVITY OF THE LATER.
 RX MEDLINE: 90368561.

RA Webb R., Reddy K.J., Sherman L.A.;
 RI "Regulation and sequence of the Synchococcus sp. strain PCC 7942
 RI groESL operon, encoding a cyanobacterial chaperonin."

RC J. Bacteriol. 172:5079-5080(1990).
 RN [1]

RP THE ATPASE ACTIVITY OF THE LATER.
 RX MEDLINE: 90368561.

RA Webb R., Reddy K.J., Sherman L.A.;
 RI "Regulation and sequence of the Synchococcus sp. strain PCC 7942
 RI groESL operon, encoding a cyanobacterial chaperonin."

RC J. Bacteriol. 172:5079-5080(1990).
 RN [1]


```

DR EMBL: U30821; AAA81293.1;
DR PRINTS: PR00297; CHAPERONIN10
DR PROSITE: PS00681; CHAPERONINS_CPN10; 1.
DR PFAM: PF00186; cpn10; 1.
DR Chaperone; Cyanelle.
SQ
SEQUENCE 103 AA; 11240 MW; FF768479C0A90520 CRC64;

Query Match 34.4%; Score 78; DB 1; Length 103.
Best Local Similarity 50.0%; Pred. No. 3,44e-01;
Matches 8; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

DD 99 EYVLLSEKDLAIKA 103
      ||||| |||||
QY 1 EYVLLSARDVLAWS 16

RESULT 14
ID NTPO_ENTHR STANDARD PRI 230 AA.
AC P43435
DI 01-NOV-1995 (Rel. 32, Created)
DI 01-NOV-1995 (Rel. 32, Last sequence update)
DI 01-OCT-1996 (Rel. 34, Last annotation update)
DE V-TYPE SODIUM ATP SYNTHASE SUBUNIT D (EC 3.6.1.34) (NA(+)-
TRANSLATING ATPASE SUBUNIT D).
GN NTPS.
OS Enterococcus hirae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;
OC Enterococcus.
RN (1)
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-14.
RC STRAIN=ATCC 9790;
RX MEDLINE: 94205259.
RA Takase K., Kakinuma S., Yamato I., Konishi K., Igarashi K.,
RA Kakinuma Y.
RT vacuolar-type Na(+)-translocating ATPase of Enterococcus hirae.
CC J. Biol. Chem. 269:11037-11044(1994).
CC -- FUNCTION: INVOLVED IN ATP-DRIVEN SODIUM EXTRUSION.
CC -- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS V-TYPE SUBUNIT D.
CC V-TYPE SODIUM ATPASE NTPS AND ARCHEAL ATPASE SUBUNIT D.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL database at
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sb.ch).
CC
CC EMBL: D17462; BA84472.1;
CC Hydrolyase; Hydrolase for transport.
SQ
SEQUENCE 230 AA; 27193 MW; 950421364E56EF09 CRC64;

Query Match 34.4%; Score 78; DB 1; Length 230.
Best Local Similarity 40.0%; Pred. No. 3,44e-01;
Matches 12; Conservative 7; Mismatches 10; Indels 1; Gaps 1;

DB 78 DELLALPAENVSISVVEKNIMSVKVPIMNF 107
      ||||| ||||| |||||
QY 1 EYVLLSARDV-LAWSRRKMKVIVAFNOF 29

RESULT 15
ID CHICLLEPIN STANDARD PRI: 96 AA.
AC P35472;
DI 01-JUN-1994 (Rel. 29, Created)
DI 01-JUN-1994 (Rel. 29, Last sequence update)
DI 15-DEC-1998 (Rel. 37, Last annotation update)
DE 10 KD CHAPERONIN (PROTEIN CPN10) (PROTEIN GROES) (HEAT SHOCK 10 KD
PROTEIN).
GN GROES OR HSP10.
OS Leptospira interrogans.
OC Bacteria; Spirochaetales; Leptospiraceae; Leptospira.

```

```

(1)
SEQUENCE FROM N.A.
STRAIN-SEROVAR COPENHAGEN / WJUNBERG:
MEDLINE: 9313752.
Ballard S.A., Seders R.P., Breutink-Piym N.M., Eylar J.A.M.,
Faine S., Adler B.:
"molecular analysis of the hsp (groE) operon of Leptospira
interrogans serovar Copenhageni";
Mol. Microbiol. 8:719-751(1993).
(2)
SEQUENCE FROM N.A.
STRAIN-SEROVAR LAI / HN-1:
Kim M.J., And S.Y.:
Submitted (Nov-1997) to the EMBL/Genbank/DBJ databases.
-1- FUNCTION: BINDS TO CPN50 IN THE PRESENCE OF MG-ATP AND SUPPRESSES
THE ATPASE ACTIVITY OF THE LATTER.
-1- SUBUNIT: HEPTAMER OF 7 SUBUNITS ARRANGED IN A RING
BY SIMILARITY.
-1- INDUCTION: BY HEAT SHOCK.
-1- SIMILARITY: BELONGS TO THE GROES CHAPERONIN FAMILY.
This SWISS-PRO entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL: U14682; AAA71991.1;
EMBL: AF022910; AAR86964.1;
PIR: S34937; S34937.
PRINTS: PR00297; CHAPERONIN10
PROSITE: PS01681; CHAPERONINS_CPN10; 1.
PFAM: PF00166; cpn10; 1.
Chaperone; Heat Shock.
SEQUENCE 96 AA: 10562 MW: 528508079D8F3467 CRC64:
Query Match 33.5% Score 76; DB 1; Length 96;
Best Local Similarity 62.5% Pred. No. 7 326-01;
Matches 10; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
LF 8: EYDITRESDIAYVKK 96
10:13 11:11 1
DY 2 EYDILSAPDVAVVSK 17
Search completed: Mon Jun 19 16:11:06 2000
Total time 1.17 secs.

```

WATERFALL
(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit,
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MParch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Mon Jun 19 16:11:24 2000: MasPar time 9.82 seconds
Tabular output not generated. 225,934 Million cell updates/sec

Title: SUS-09-142-524A-10
Description: (1-32) from US09142524A.pep
Perfect Score: 227
Sequence: 1 EYLILSARDVLAVYSKRMKV-VAFNQEGPN 32

Scoring table: PAM 150
Gap 11

Searched: 225678 seqs, 59334122 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: strebl12
1:sp_archaea 2:sp_bacteria 3:sp_fungi 4:sp_human
5:sp_invertebrate 6:sp_mammal 7:sp_mhc 8:sp_organelle
9:sp_phase 10:sp_plant 11:sp_rodent 12:sp_unclassified
13:sp_vertebrate 14:sp_virus

Statistics: Mean 33.599; Variance 54.501; scale 0.616

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description | Pred. No. |
|------------|-------|-------------|--------|-------|-------------------------------|-----------|
| 1 | 114 | 50.2 | 100 | 2 | 086017 10 KD CHAPERONIN (PROT | 2.33e-07 |
| 2 | 98 | 43.2 | 357 | 10 | Q92NU7 POLLEN MAJOR ALLERGEN | 2.42e-04 |
| 3 | 98 | 43.2 | 357 | 10 | Q92NU7 CHAGI PRECURSOR. | 2.42e-04 |
| 4 | 97 | 42.7 | 459 | 10 | Q23565 PUTATIVE PEPTIDE LYASE | 3.68e-04 |
| 5 | 92 | 40.5 | 354 | 10 | Q23565 PUTATIVE PEPTIDE LYASE | 2.90e-03 |
| 6 | 91 | 40.1 | 354 | 10 | Q65456 PEPTIDE LYASE LIKE PRO | 4.35e-03 |
| 7 | 88 | 38.8 | 242 | 5 | Q96136 ERCC1-LIKE EXCISION RE | 1.45e-02 |
| 8 | 88 | 38.8 | 390 | 10 | Q65488 F12F1.22 PROTEIN | 1.45e-02 |
| 9 | 84 | 37.0 | 455 | 10 | Q65450 PUTATIVE PEPTIDE LYASE | 7.02e-02 |
| 10 | 80 | 35.2 | 310 | 8 | Q92722 NADH DEHYDROGENASE SUB | 3.25e-01 |
| 11 | 80 | 35.2 | 328 | 8 | Q37529 NADH DEHYDROGENASE (UB | 3.25e-01 |
| 12 | 79 | 34.8 | 398 | 10 | Q43783 PEPTIDE LYASE (EC 4.2 | 4.73e-01 |
| 13 | 78 | 34.4 | 330 | 8 | Q922Q3 NADH-UBIQUINONE OXIDOR | 6.88e-01 |
| 14 | 77 | 33.9 | 305 | 2 | Q9X716 PUTATIVE OXIDOREDUCTAS | 9.96e-01 |
| 15 | 77 | 33.9 | 768 | 5 | P92890 F55H12.1 PROTEIN | 9.96e-01 |
| 16 | 76 | 33.5 | 103 | 2 | Q9X2U6 10 KD CHAPERONIN (PROT | 1.44e-00 |
| 17 | 76 | 33.5 | 323 | 10 | Q42576 MITOCHONDRIAL NADH DEH | 1.44e-00 |
| 18 | 75 | 33.5 | 325 | 8 | P92558 NADH DEHYDROGENASE SUB | 1.44e-00 |
| 19 | 75 | 33.5 | 331 | 8 | Q37565 NADH DEHYDROGENASE (UB | 1.44e-00 |
| 20 | 75 | 33.0 | 96 | 2 | O51384 HYPOTHETICAL 11.4 KD P | 2.07e-00 |

| | | | | | | |
|----|----|------|------|----|--------------------------------|----------|
| 21 | 75 | 33.0 | 225 | 1 | Q98715 225AA LONG HYPOTHETICA | 2.37e-00 |
| 22 | 75 | 33.0 | 4513 | 2 | Q92G15 TYPE I POLYPEPTIDE SYNT | 2.37e-00 |
| 23 | 74 | 32.6 | 302 | 5 | Q17137 F31F4.3 PROTEIN. | 2.97e-00 |
| 24 | 74 | 32.6 | 1006 | 4 | C15197 EPH-FAMILY RECEPTOR FR | 2.97e-00 |
| 25 | 73 | 32.2 | 88 | 2 | Q5C304 GROES (FRAGMENT) | 4.24e-00 |
| 26 | 73 | 32.2 | 129 | 8 | Q95867 NADH-UBIQUINONE OXIDOR | 4.24e-00 |
| 27 | 73 | 32.2 | 226 | 10 | Q15667 PUTATIVE PEPTIDE LYASE | 4.24e-00 |
| 28 | 73 | 32.2 | 401 | 10 | Q15667 PEPTIDE LYASE PRECURSOR | 4.24e-00 |
| 29 | 73 | 32.2 | 517 | 10 | Q8C729 PUTATIVE CYTOCHROME P4 | 4.24e-00 |
| 30 | 73 | 32.2 | 733 | 5 | Q22561 119H10.8 PROTEIN | 4.24e-00 |
| 31 | 73 | 32.2 | 841 | 2 | Q95709 RELA PROTEIN | 4.24e-00 |
| 32 | 73 | 32.2 | 2823 | 5 | Q45614 122A3.6 PROTEIN (FRAGM | 4.24e-00 |
| 33 | 72 | 31.7 | 301 | 2 | Q53358 OXIDOREDUCTASE | 6.05e-00 |
| 34 | 72 | 31.7 | 374 | 2 | P23273 CAP5G. | 6.05e-00 |
| 35 | 72 | 31.7 | 374 | 2 | P23273 CAP5G. | 6.05e-00 |
| 36 | 72 | 31.7 | 438 | 10 | Q43862 POLLEN SPECIFIC PEPTAT | 6.05e-00 |
| 37 | 72 | 31.7 | 1014 | 11 | Q8B644 EPH/ELK RECEPTOR LIKE | 4.05e-00 |
| 38 | 71 | 31.3 | 369 | 13 | Q91790 DUAL SPECIFICITY PROTE | 6.59e-00 |
| 39 | 71 | 31.3 | 405 | 10 | Q24416 PEPTIDE LYASE | 8.59e-00 |
| 40 | 71 | 31.3 | 450 | 10 | Q40319 PEPTIDE LYASE HOMOLOG | 8.59e-00 |
| 41 | 71 | 31.3 | 465 | 4 | Q94889 KIAA0795 PROTEIN (FRAG | 8.59e-00 |
| 42 | 70 | 30.8 | 482 | 5 | Q22329 COSMID T3H8. | 1.22e-01 |
| 43 | 70 | 30.8 | 506 | 14 | G11997 GLYCOPROTEIN PRECURSOR | 1.22e-01 |
| 44 | 70 | 30.8 | 509 | 14 | G11997 GLYCOPROTEIN PRECURSOR | 1.22e-01 |
| 45 | 70 | 30.8 | 1106 | 13 | P79726 RECEPTOR TYROSINE KINA | 1.22e-01 |

ALIGNMENTS

RESULT 1
ID 086017 PRELIMINARY: PRI: 100 AA.
AC 086017:
DT 01-NOV-1998 (Tremblrel, 08, Created)
DI 01-NOV-1998 (Tremblrel, 08, Last sequence update)
DI 01-NOV-1998 (Tremblrel, 12, Last annotation update)
DE 10 KD CHAPERONIN (PROTEIN C.N.0) (PROTEIN GROES).
GN GROES.
OS Mycobacterium avium, and Mycobacterium paratuberculosis
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacti Neae; Mycobacteriaceae; Mycobacterium.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-485 TYPE 21.
RA CRET R., PIETROBONO R., F., KINI L., OREFICI G.:
RT "Sequence of the GroES1 protein of Mycobacterium avium comprising the
RT gene encoding the cpn19 protein and a portion of the gene encoding the
RT cpn60-1 protein."
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES-8M PARATUBERCULON STRAIN-ATCC 19658, ATCC 25241
RA GORB A.J., PRUTHINGHAM R.:
RT "The GroES antigens of Mycobacterium avium and Mycobacterium
RT paratuberculosis."
RL Microbiol. 0:0-0(1999).
CC -- FUNCTION: BINDS TO CPN60 IN THE PRESENCE OF MG-ATP AND SUPPRESS
CC THE ATPASE ACTIVITY OF THE LATTER (BY SIMILARITY).
CC -- SUBUNIT: HEPTAMER OF 7 SUBUNITS ARRANGED IN A RING (BY
CC SIMILARITY).
CC -- SIMILARITY: BELONGS TO THE GROES CHAPERONIN FAMILY.
DR EMBL: AF079544; AAC3192.1;
DR EMBL: AF071829; AAD23277.1;
DR EMBL: AF071828; AAD23276.1;
DR PFAM: PF00166; cpn10; 1.
DR PRINTS: PR00297; CHAPERON.N10.
KW Chaperone; Heat shock.
SQ SEQUENCE 100 AA; 10748 MW; 0C4954DA CRC32;

Query Match 50.2%; Score 114; DB 2; Length 100;
Best Local Similarity 100.0%; Pred. No. 2.33e-07;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Caps 0;

Db 84 EYLLISARDVLAVWSK 100
 QY 1 EYLLISARDVLAVWSK 17

RESULT 2
 ID Q92N07 PRELIMINARY: PRT: 367 AA.
 AC Q92N07
 DT 01-MAY-1999 (TEMBLrel. 10, Created)
 DT 01-MAY-1999 (TEMBLrel. 10, Last sequence update)
 DT 01-NOV-1999 (TEMBLrel. 12, Last annotation update)
 DE POLLEN MAJOR ALLEGEN 1-2
 CS Juniperus ashei (Ozark white cedar)
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC e-phylllophytes; Spermatophyta; Coniferales;
 CC Taxodiaceae; Juniperus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MIDOR-HORIUT: I.M., GOLDBLUM R.M., KUKOSKY A., WOOD T.G.,
 RA BROOKS E.G.
 RT "Molecular cloning of mountain cedar (Juniperus ashei) pollen major
 RT allergen, Jun a 1."
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF066563; RAD03609.1;
 DR EMBL: AF066562; RAD03608.1;
 DR EMBL: AF066561; RAD03607.1;
 DR MENDEL: 36544; Junas:1088;36544.
 DR MENDEL: 36545; Junas:1088;36545.
 KW PRINTS: PR00807; AMBALLERG.
 SQ SEQUENCE 367 AA: 39824 MW: 4C2DB630 CRC32:

Query Match 43.2% Score 98: DB 10: Length 367;
 Best Local Similarity 86.7% Pred. No. 2,42e-04;
 Matches 13: Conservative 1: Mismatches 1: Indels 0: Gaps 0:

Db 232 KSMKVTVAFNQFGPN 246
 QY 18 RRMKVTVAFNQFGPN 32

RESULT 3
 ID Q96385 PRELIMINARY: PRT: 376 AA.
 AC Q96385
 DT 01-FEB-1997 (TEMBLrel. 02, Created)
 DT 01-FEB-1997 (TEMBLrel. 02, Last sequence update)
 DT 01-NOV-1999 (TEMBLrel. 12, Last annotation update)
 DE CHAOI FRECURSOR
 CS Chamaecyparis obtusa.
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC e-phylllophytes; Spermatophyta; Coniferales;
 CC Taxodiaceae; Chamaecyparis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA SUZUKI M., KOMIYAMA N., ITOH M., ITOH H., SONE T., KUNO K., TAKAGI I.,
 RA OHTA N.
 RT "Purification, characterization and molecular cloning of Chamaecyparis
 RT major allergen of Chamaecyparis obtusa (Japanese cypress) pollen."
 RL Mol. Immunol. 33:451-460(1996).
 DR EMBL: E45404; BAA08246.1;
 DR MENDEL: 7626; Chaoi:1088;7626.
 DR PFAM: PF00544; pec_lyase.1
 DR PRINTS: PR00807; AMBALLERG.
 KW Signal
 FT SIGNAL 1 21 POTENTIAL.
 FT CHAIN 22 375 CHAOI.
 SQ SEQUENCE 375 AA: 40258 MW: A098452 CRC32:

Query Match 43.2% Score 98: DB 10: Length 375;
 Best Local Similarity 86.7% Pred. No. 2,42e-04;
 Matches 13: Conservative 1: Mismatches 1: Indels 0: Gaps 0:

Db 232 KSMKVTVAFNQFGPN 246
 QY 18 RRMKVTVAFNQFGPN 32

QY 18 RRMKVTVAFNQFGPN 32

RESULT 4
 ID Q23665 PRELIMINARY: PRT: 459 AA.
 AC Q23665
 DT 01-JAN-1998 (TEMBLrel. 05, Created)
 DT 01-JAN-1998 (TEMBLrel. 05, Last sequence update)
 DT 01-NOV-1999 (TEMBLrel. 12, Last annotation update)
 DE PEPTIDE LYASE
 CS Arabidopsis thaliana (Mouse ear cress)
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC e-phylllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 CC core eudicots; Rosidae; euro. 11; Brassicales; Brassicaceae;
 CC Arabidopsis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 97422403.
 RA KULIKAUSSAS R., MCCORMICK S.
 RT "Identification of the tobacco and Arabidopsis homologues of the
 RT pollen-expressed LAI59 gene of tomato."
 RL Plant Mol. Biol. 34:809-814(1997).
 DR EMBL: U83613; AAB9759.1;
 DR MENDEL: 25559; Arabid:1088;25559.
 DR PFAM: PF00544; pec_lyase.1.
 DR PRINTS: PR00807; AMBALLERG.
 KW Lyase.
 SQ SEQUENCE 459 AA: 51420 MW: 41E5908 CRC32:

Query Match 42.7% Score 97: DB 10: Length 459;
 Best Local Similarity 46.7% Pred. No. 3,58e-04;
 Matches 14: Conservative 9: Mismatches 5: Indels 2: Gaps 2:

Db 298 QEVNLFQARGE-HVIDKK-MQITVAFNHF 325
 QY 1 EYLLISARDVLAVWSKRRKVTVAFNHF 33

RESULT 5
 ID Q65457 PRELIMINARY: PRT: 394 AA.
 AC Q65457
 DT 01-AUG-1998 (TEMBLrel. 07, Created)
 DT 01-AUG-1998 (TEMBLrel. 07, Last sequence update)
 DT 01-NOV-1999 (TEMBLrel. 12, Last annotation update)
 DE PEPTIDE LYASE LIKE PROTEIN
 CS Arabidopsis thaliana (Mouse ear cress)
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC e-phylllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 CC core eudicots; Rosidae; euro. 11; Brassicales; Brassicaceae;
 CC Arabidopsis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA BEVAN M., WEDLER H., WAMBUTT R., BANCROFT T., MEWES H.W., MAYER K.,
 RA SCHUELLER C.
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU ARABIDOPSIS SEQUENCING PROJECT:
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: ALG22140; CAA18112.1;
 DR MENDEL: 29068; Arabid:1088;29068.
 DR PFAM: PF00544; pec_lyase.1.
 DR PRINTS: PR00807; AMBALLERG N.
 KW Lyase.
 SQ SEQUENCE 394 AA: 43299 MW: 889A93C9 CRC32:

Query Match 40.5% Score 92: DB 10: Length 394;
 Best Local Similarity 70.6% Pred. No. 2,90e-03;
 Matches 12: Conservative 2: Mismatches 3: Indels 0: Gaps 0:

Db 251 VKDVKMKVTVAFNHF 267

QY 15 VSKRMKVTVAFNFGSP 31
 :|||||:|||||

RESULT 6
 ID O65456 PRELIMINARY: PRT: 394 AA.
 AC O65456
 DT 01-AUG-1998 (TREMblrel. 07, Created)
 DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)
 DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)
 DE F12F1.22 PROTEIN.
 GN F12F1.22
 OS PECTATE LYASE LIKE PROTEIN.
 OS Arabidopsis thaliana (Mouse-ear cross).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 OC core eudicots; Rosidae; eustosids II; Brassicales; Brassicaceae;
 OC Arabidopsis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA BEVAN M., WEDLER H., WAMBATT R., BANCROFT I., MEWES H.W., MAYER K.,
 PA SCHUELLER C.
 RI Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU ARABIDOPSIS SEQUENCING PROJECT;
 RI Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: A022140; CAAL9111.1;
 DR MENDEL: 29747; Arab:1085;29747.
 DR PFAM: PF00544; pec_lyase; 1.
 DR PRINTS: PRO0807; AMBALLERGEN.
 KW Lyase.
 SQ SEQUENCE 394 AA: 43476 MW: 95399178 CRC32;
 Query Match 40.8% Score 91; DB 10; Length 394;
 Best Local Similarity 84.6% Pred. No. 4.35e-03;
 Matches 11; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

DB 255 KMKVTVAFNFGSP 267
 :|||||:|||||
 QY 19 KMKVTVAFNFGSP 31

RESULT 7
 ID O56136 PRELIMINARY: PRT: 242 AA.
 AC O56136
 DT 01-MAY-1998 (TREMblrel. 10, Created)
 DT 01-MAY-1998 (TREMblrel. 10, Last sequence update)
 DT 01-MAY-1999 (TREMblrel. 10, Last annotation update)
 DE ERCC1-LIKE EXCISION REPAIR PROTEIN.
 GN PF06160W.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA GARDNER M.J., TEITELIN H., CARUCCI D.J., CUMMINGS L.M., ARAVIND L.,
 RA KOONIN E.V., SHALLOM S., MASON T., YU K., FUJII C., PEPERSON J.,
 RA SHEN K., JING J., ASTON C., LAI Z., SCHWARTZ D.C., PERTEA M.,
 RA SALZBERG S., ZHOU S., SUTTON G.S., CLAYTON R., WHITE C., SMITH H.O.,
 RA FRASER C.M., ADAMS M.D., VENTER J.C., HOFFMAN S.L.;
 RT *Chromosome 2 sequence of the human malaria parasite Plasmodium
 falciparum.
 RL Science 282:1126-1132(1998).
 DR EMBL: AE001377; AAC71822.1;
 SQ SEQUENCE 242 AA: 28287 MW: 4590AD3C CRC32;

Query Match 38.8% Score 88; DB 5; Length 242;
 Best Local Similarity 40.6% Pred. No. 1.45e-02;
 Matches 13; Conservative 8; Mismatches 11; Indels 0; Gaps 0;
 DB 45 EYLLSARDVLAIVVSKRMKVTVAFNFGSP 76
 :|||||:|||||
 QY 1 EYLLSARDVLAIVVSKRMKVTVAFNFGSP 32

RESULT 8
 ID O65388 PRELIMINARY: PRT: 390 AA.
 AC O65388
 DT 01-AUG-1998 (TREMblrel. 07, Created)
 DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)
 DT 01-NOV-1999 (TREMblrel. 12, Last annotation update)
 DE F12F1.22 PROTEIN.
 GN F12F1.22
 OS Arabidopsis thaliana (Mouse-ear cross).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 OC core eudicots; Rosidae; eustosids II; Brassicales; Brassicaceae;
 OC Arabidopsis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA VISOISKATA V.S., OSBORNE B.L., SCHWARTZ J.R., TORIEMI M., YU C.,
 RA KWAN A., OUI O., LIU S., BUEHLER E., CONWAY A.B., CONWAY A.R.,
 RA DEWAR K., FENG C., KIM C., KURTZ D., PALM C.J., LI Y., SHINN P.,
 RA SUN H., DAVIS R.W., ECKER J.R., FEDERSPIEL N.A., THEODOSIS A.,
 RI Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC002131; AAC17625.1;
 DR MENDEL: 29024; Arab:1088;29024.
 DR PFAM: PF00544; pec_lyase; 1.
 DR PRINTS: PRO0807; AMBALLERGEN.
 SQ SEQUENCE 390 AA: 43354 MW: E0F12FB4 CRC32;

Query Match 38.8% Score 88; DB 10; Length 390;
 Best Local Similarity 58.8% Pred. No. 1.45e-02;
 Matches 10; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

DB 241 VEDKIMRVTVAFNFGSP 257
 :|||||:|||||
 QY 15 VSKRMKVTVAFNFGSP 31

RESULT 9
 ID O64510 PRELIMINARY: PRT: 455 AA.
 AC O64510
 DT 01-AUG-1998 (TREMblrel. 07, Created)
 DT 01-AUG-1998 (TREMblrel. 07, Last sequence update)
 DT 01-MAY-1999 (TREMblrel. 12, Last annotation update)
 DE PUTATIVE PEPTIDE LYASE.
 GN F12F1.22
 OS Arabidopsis thaliana (Mouse-ear cross).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 OC core eudicots; Rosidae; eustosids II; Brassicales; Brassicaceae;
 OC Arabidopsis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN-CV, COLUMBIA;
 RA ROUSLEY S.D., LIN X., KETCHUM K.A., CROSBY M.L., BRANDON R.C.,
 RA SYKES S.M., KAUL S., MASON J.M., KERLAVAGE A.R., ADAMS M.D.,
 RA SOMERVILLE C.R., VENTER J.C.;
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC002521; AAC03530.1;
 DR MENDEL: 28505; Arab:1088;28505.
 DR PFAM: PF00544; pec_lyase; 1.
 DR PRINTS: PRO0807; AMBALLERGEN.
 KW Lyase.
 SQ SEQUENCE 455 AA: 51257 MW: EBA0B82D CRC32;

Query Match 37.0% Score 84; DB 10; Length 455;
 Best Local Similarity 36.7% Pred. No. 7.02e-02;
 Matches 11; Conservative 11; Mismatches 6; Indels 2; Gaps 2;
 DB 294 DEVMLFGGTNK-DVIDKK-MQIVAFNFGSP 321
 :|||||:|||||
 QY 1 EYLLSARDVLAIVVSKRMKVTVAFNFGSP 30

```

RESULT 10
ID Q92Y22 PRELIMINARY: PRI: 310 AA
AC Q92Y22
DT 01-MAY-1999 (TRENBLREL: 01, Created)
DT 01-MAY-1999 (TRENBLREL: 10, Last sequence update)
DT 01-MAY-1999 (TRENBLREL: 12, Last annotation update)
CE NADH DEHYDROGENASE SUBUNIT 1 (EC 1.6.5.3)
GN NADH
OS Pedinomonas minor
OC Mitochondrion
CC Eukaryota; Viridiplantae; Chlorophyta; Pedinophyceae; Pedinomonadales;
PC Pedinomonadales; Pedinomonas
AN [1]
RP SEQUENCE FROM N.A.
RC STRAIN: UTEX LP 1350
RA TURNER M., LEMIEUX C., BURGER G., LANG B F., OTIS C., PLANTE L.,
RA GRAY M.W.
RT "The complete mitochondrial DNA sequences of Neptodermis olivacea and
RT Pedinomonas minor: two radically different evolutionary patterns
RT within the green algae."
RC Submitted (EC-1998)
CC EMBL: AF116775; AAD19672.1
CC OXPHOSOMES; Mitochondrion
DR PFAM: P00146; NADHdh: 1
KW Ubiquinone: Oxidoreductase; Mitochondrion; NAD: Transmembrane
SQ SEQUENCE 310 AA: 34938 MW: 27624CA CRC32:

Query Match 35.2% Score 80: DB 8: Length 310:
Best Local Similarity 31.7% Pred. No. 3,25e-01:
Matches 10: Conservative 6: Mismatches 8: Indels 0: Gaps 0:

DB 28 FLVLAERKVLASQKRGKGNVGVGIFGP 51
QY 3 YLLSARDVLAVVSKRRKRVVAFNQFQF 26

RESULT 11
ID Q37629 PRELIMINARY: PRI: 328 AA
AC Q37629
DT 01-NOV-1996 (TRENBLREL: 01, Created)
DT 01-NOV-1996 (TRENBLREL: 10, Last sequence update)
DT 01-NOV-1999 (TRENBLREL: 12, Last annotation update)
CE NADH DEHYDROGENASE (UBIQUINONE), SUBUNIT 1 (EC 1.6.5.3)
GN NADH
OS Protobotheca wickerhamii
OC Mitochondrion
CC Eukaryota; Viridiplantae; Chlorophyta; Trebouxiophyceae; Chlorellales;
PC Chlorellaceae; Protobotheca
AN [1]
RP SEQUENCE FROM N.A.
RC STRAIN: 263-11
RA WOLFE G., PLANTE L., LANG B.F., KUECK G., BURGER G.
RT "Complete sequence of the mitochondrial DNA of the chlorophyte alga
RT Protobotheca wickerhamii: Gene content and genome organization."
RL J. Mol. Biol. 237:75-86(1994)
DR EMBL: U02970; AAO12657.1
DR WENSEL: 21568; Protoboth: 21568
DR PFAM: P00146; NADHdh: 1
KW Ubiquinone: Oxidoreductase; Mitochondrion; NAD: Transmembrane
SQ SEQUENCE 328 AA: 36437 MW: F1269545 CRC32:

Query Match 35.2% Score 80: DB 8: Length 328:
Best Local Similarity 33.3% Pred. No. 3,25e-01:
Matches 10: Conservative 8: Mismatches 11: Indels 1: Gaps 1:

DB 25 FLVLAERKVLASQKRGKGNVGVGIFGP 54
QY 3 YLLSARDVLAVVSKRRKRVVAFNQFQF 31

RESULT 12
ID Q43783 PRELIMINARY: PRI: 338 AA
AC Q43783

```

```

DT 01-NOV-1996 (TRENBLREL: 01, Created)
DT 01-NOV-1996 (TRENBLREL: 01, Last sequence update)
DT 01-NOV-1999 (TRENBLREL: 12, Last annotation update)
DE PECTATE LYASE (EC 4.2.2.2) (PECTATE TRANSPELIMINASE) (FRAGMENT)
GN PEL
OS Musa acuminata (Banana)
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta
OC euphyllophytes; Spermatophyta; Magnoliophyta; Liliopsida
OC Zingiberales; Musaceae; Musa
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN: CV, DWARF CAVENTISH; TISSUE-RIPE FRUIT
RA DOMINGUEZ-PUCOJANER E.
RC Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN: CV, DWARF CAVENTISH; TISSUE-RIPE FRUIT
RA DOMINGUEZ-PUCOJANER E.
RC Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC "CATALYTIC ACTIVITY: ELIMINATIVE CLEAVAGE OF PECTATE TO GIVE
CC OCTOSACCHARIDES WITH 4-DEOXY-ALPHA-D-GLUCO-4-ENOLIN SYNOMONAS AI
CC THEIR NON-REDUCING ENDS."
CC EMBL: X52943; CAA63496.1
CC MENDEL: 9136; Musac: 1088.9.36.
DR PFAM: PFC0544; PectLyase: 1
KW Lyase
FT NON-TER
SQ SEQUENCE 398 AA: 43712 MW: E5E3B38A CRC32:

Query Match 34.8% Score 75: DB 10: Length 398:
Best Local Similarity 41.4% Pred. No. 4,73e-01:
Matches 12: Conservative 8: Mismatches 7: Indels 2: Gaps 1:

DB 246 MLGHSDELTS-SKSMQVTFAPNHFGED 272
QY 4 YLLSARDVLAVVSKRRKRVVAFNQFQF 52

RESULT 13
ID Q92YJ3 PRELIMINARY: PRI: 330 AA
AC Q92YJ3
DT 01-MAY-1999 (TRENBLREL: 10, Created)
DT 01-MAY-1999 (TRENBLREL: 10, Last sequence update)
DT 01-NOV-1999 (TRENBLREL: 12, Last annotation update)
DE NADH-UBIQUINONE OXIDOREDUCTASE CHAIN 1 (EC 1.6.5.3)
GN NADH
OS Cyanidioschyzon merolae
OC Mitochondrion
CC Eukaryota; Rhodophyta; Rhodospirillaceae; Rhodospirillum rubrum
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN: N100
RX MEDLINE: 99030526
RA OHYA N., SATO N., KUROIWA T.
RT "Structure and organization of the mitochondrial genome of the
RT unicellular red alga Cyanidioschyzon merolae deduced from the complete
RT nucleotide sequence."
RL Nucleic Acids Res 24:5190-5198(1998)
DR EMBL: D89963; BAA36521.1
DR PROSITE: PS00667; COMPLEX1_NCL_1: 1
DR PROSITE: PS00668; COMPLEX1_NCL_2: 1
KW Ubiquinone: Oxidoreductase; Mitochondrion; NAD: Transmembrane
SQ SEQUENCE 330 AA: 36961 MW: 5B084D2E CRC32:

Query Match 34.4% Score 78: DB 8: Length 330:
Best Local Similarity 33.3% Pred. No. 6,88e-01:
Matches 10: Conservative 8: Mismatches 11: Indels 1: Gaps 1:

DB 26 YLLAERKILGATCKKRGKGNVGVGIFGP 55
QY 3 YLLSARDVLAVVSKRRKRVVAFNQFQF 31

```

```

RESULT 14
ID Q9X7T6 PRELIMINARY: PRT: 305 AA.
AC Q9X7T6
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DI 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DE 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE PUTATIVE OXIDOREDUCTASE.
GN SC5H1.28C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2)
RA OLIVER K., HARRIS D.
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome."
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2)
RA JAMES K.D., PARKHILL J., BARRELL B.G., PAJANDREAM M.A.
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2)
RA MEDLINE: 9700035.
RA REDENBACH M., KIESER H.M., DENAPATE D., EICHNER A., CULLUM J.,
RA KINASHI H., HOPWOOD D.A.
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome."
RL Microbiol. 21:77-96(1996).
CR EMBL: AL049863; CAB42953.1;
SC SEQUENCE 305 AA: 32531 MW: 54145AB9 CRC32:

Query Match 33.9% Score 77: DB 2: Length 305;
Best Local Similarity 41.9% Pred. No. 9,96e-01;
Matches 13: Conservative 5: Mismatches 13: Indels 0: Gaps 0:

LD 127 EENLRQGRDHLDDVYLRMRQDSVAEFGA 157
      | | | | | | | | | | | | | | | |
QY 1 REYLILSARDVLAVYVKRKRKMKVIVAFNFGP 31

RESULT 15
ID P90890 PRELIMINARY: PRT: 768 AA.
AC P90890
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DI 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DE 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE F58H12.1 PROTEIN.
GN F58H12.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;
OC Rhabditina; Rhabditoidea; Rhabditidae; Pelodierinae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RA DOBSON R.
RC Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE: 94150718.
RA WILSON R., AINSCOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA CRATON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIEKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.

```

```

RT elegans.
RL Nature 368:32-38(1994).
DR EMBL: Z81091; CAB03142.1;
DR PFAM: PF00209; SNF: 2.
DR PRINTS: PRO0176; NANEUSM20PT.
SQ SEQUENCE 768 AA: 85132 MW: D25DRG32 CRC32:

Query Match 33.9% Score 77: DB 5: Length 768;
Best Local Similarity 27.3% Pred. No. 9,96e-01;
Matches 6: Conservative 11: Mismatches 5: Indels 0: Gaps 0:

DB 669 FLVIPGFAIAAVINTRRMNPI 689
      | | | | | | | | | | | | | | | |
QY 3 YLIIISARDVLAVYVKRKRKMKV 24

Search completed: Mon Jun 19 16:11:45 2000
Job time : 21 secs.

```

| Result
No. | Score | Query | | Length | DB | ID | Description | Pred. No. |
|---------------|-------|-------|-------|--------|---------|-------------------------|-------------|-----------|
| | | Match | Count | | | | | |
| 1 | 143 | 100.0 | 20 | 1 | W42130 | T-cell epitope peptide | 3.71e-08 | |
| 2 | 143 | 100.0 | 3 | 1 | W27372 | Multi-epitope peptide | 3.71e-08 | |
| 3 | 143 | 100.0 | 354 | 1 | W04344 | Chamaecyparis obtusa p | 3.71e-08 | |
| 4 | 143 | 100.0 | 354 | 1 | W4121 | Japanese cypress pollen | 3.71e-08 | |
| 5 | 143 | 100.0 | 375 | 1 | W04345 | Chamaecyparis obtusa p | 3.71e-08 | |
| 6 | 123 | 86.0 | 367 | 1 | R45577 | Jun s i | 8.16e-06 | |
| 7 | 123 | 86.0 | 370 | 1 | R45578 | Jun v i | 8.16e-06 | |
| 8 | 143 | 79.0 | 20 | 1 | R82498 | Cry j i Japanese Cedar | 1.15e-04 | |
| 9 | 143 | 79.0 | 20 | 1 | R45549 | Cry j i pollen allergen | 1.15e-04 | |
| 10 | 113 | 79.0 | 30 | 1 | W44582 | T-cell epitope peptide | 1.15e-04 | |
| 11 | 113 | 79.0 | 34 | 1 | W27371 | Multi-epitope peptide | 1.15e-04 | |
| 12 | 113 | 79.0 | 353 | 1 | R73388 | Japanese cedar pollen | 1.15e-04 | |
| 13 | 113 | 79.0 | 353 | 1 | R83587 | Cedar pollen allergen | 1.15e-04 | |
| 14 | 113 | 79.0 | 374 | 1 | R31937 | Cry j i | 1.15e-04 | |
| 15 | 113 | 79.0 | 374 | 1 | R45541 | Cry j i pollen allergen | 1.15e-04 | |
| 16 | 113 | 79.0 | 374 | 1 | R82490 | Cry j i Japanese Cedar | 1.15e-04 | |
| 17 | 113 | 79.0 | 374 | 1 | R651166 | Japanese cedar pollen | 1.15e-04 | |
| 18 | 74 | 51.7 | 20 | 1 | W42131 | T-cell epitope peptide | 2.26e+00 | |
| 19 | 72 | 50.3 | 195 | 1 | W27370 | Multi-epitope peptide | 3.65e+00 | |
| 20 | 69 | 45.3 | 123 | 1 | W41219 | T-cell epitope peptide | 7.44e+00 | |
| 21 | 67 | 42.9 | 183 | 1 | W38604 | S. pneumoniae hemolyti | 1.19e+01 | |
| 22 | 64 | 44.8 | 583 | 1 | R07347 | Amh a II/Antigen E knc | 2.39e+01 | |
| 23 | 64 | 44.8 | 583 | 1 | R07346 | Amh a II/Antigen E knc | 2.39e+01 | |

PR 10-MAR-1997; JG0740.
 PR 10-MAR-1996; JP-080702.
 PA (MEIP) MEIJI MILK PROD CO LTD.
 PI Dairiki K, Iwano A, Kino K, Kume A, Sone T;
 DR WPI: 97-470655/43.
 PT Peptide immunotherapeutic agent to treat allergic diseases -
 P1 contains multi-epitope peptide containing T cell epitope regions
 from different allergens
 PS Claim 9; Page 12; 58pp; Japanese.
 CC The present sequence represents a multi-epitope peptide which is used as
 a new immunotherapeutic agent. It comprises T cell epitope regions from 2
 or more different allergens (preferably linked via arginine or lysine
 dimers), where the T cell epitope regions have a positivity index
 greater than 10% as measured in a patient group responding to the
 allergen; have at least 70% reactivity with lymphocytes from patients
 responding to the allergen; and are not reactive with immunoglobulin E
 (IgE) antibodies from patients responsive to the allergen. The agent can
 be used to prevent and treat a wide variety of allergic diseases, e.g. by
 desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
 CC Sequence 31 AA;
 SQ

Query Match 100.0%; Score 143; DB 1; Length 31;
 Best Local Similarity 100.0%; P.ed. No. 3.71e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 IFSKNLNKLNMPLYIAGNK 20
 QY 1 IFSKNLNKLNMPLYIAGNK 20

RESULT 1
 ID W04344 standard; Protein: 354 AA.
 AC W04344.

DI 28-NOV-1996 (first entry)
 DE Chamaecyparis obtusa pollen allergen Cha o 1 (A).
 KW Pollen allergen: Cha o 1; T-cell epitope; prevention; treatment;
 KW pollinosis.
 OS Chamaecyparis obtusa.
 PN J08176192-A.
 PD 09-JUL-1996.
 PF 21-DEC-1994; 335089.
 PR 21-DEC-1994; JP-335089.
 PA (MEIP) MEIJI MILK PROD CO LTD.
 DR WPI: 95-368225/37.
 CP N-BSDB: T38519.
 PT DNA encoding chamaecyparis obtusa pollen allergen: T cell
 epitope(s) of which are useful in development of preventative and
 treating agent for C. obtusa pollen pollinosis
 PS Claim 2; Pages 15-17; 17pp; Japanese.
 CC The present sequence is the C. obtusa pollen allergen Cha o 1.
 CC The T-cell epitopes of which can be used in the development of a
 CC preventive and treating agent for C. obtusa pollen pollinosis.
 CC C. obtusa pollen (2.4 kg) was degassed with diethyl ether, and
 CC dried at room temp. overnight. Cha o 1 was sepd. from it and
 CC purified. RNA was extracted from C. obtusa pollen, and mRNA and
 CC cDNA derived.
 SQ Sequence 354 AA;

Query Match 100.0%; Score 143; DB 1; Length 354;
 Best Local Similarity 100.0%; P.ed. No. 3.71e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 71 IFSKNLNKLNMPLYIAGNK 90
 QY 1 IFSKNLNKLNMPLYIAGNK 20

RESULT 4
 ID W42121 standard; protein: 354 AA.
 AC W42121.

DI 16-JUN-1998 (first entry)
 DE Japanese cypress pollen antigen: Chaol.
 KW Japanese cypress pollen; antigen; T-cell epitope; Chaol; Chaol2;

KW diagnosis; allergy; spring tree pollen disease; pollinos.s.
 OS Chamaecyparis obtusa.
 PN W09747648-A1.
 PD 18-DEC-1997.
 PF 12-JUN-1997; J02031.
 PR 14-JUN-1996; JP-153527.
 PA (MEIP) MEIJI MILK PROD CO LTD.
 PI Dairiki K, Kino K;
 DR WPI: 98-052242/05.
 PT T-cell epitope peptide portion of Japanese cypress pollen antigens
 PR Chaol and Chaol2 used for diagnosis and treatment of spring tree
 PT pollen disease
 PS Example 1; Page 14-15; 71pp; Japanese.
 CC The present sequence represents Japanese cypress pollen antigen Chaol2.
 CC The present invention describes peptides which correspond to the T-cell
 CC epitope sites on Japanese cypress pollen antigens Chaol and Chaol2. The
 CC peptides can be used as a reagent for the diagnosis of allergy to
 CC Japanese cypress pollen, and as an antigen in the treatment and
 CC prevention of spring tree pollen disease in which the pollinosis
 CC involves reactivity to Japanese cypress pollen.
 SQ Sequence 354 AA;

Query Match 100.0%; Score 143; DB 1; Length 354;
 Best Local Similarity 100.0%; P.ed. No. 3.71e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 71 IFSKNLNKLNMPLYIAGNK 90
 QY 1 IFSKNLNKLNMPLYIAGNK 20

RESULT 5
 ID W04345 standard; Protein: 375 AA.
 AC W04345.

DI 28-NOV-1996 (first entry)
 DE Chamaecyparis obtusa pol; allergen: Cha o 1 cDNA (B).
 KW Pollen allergen: Cha o 1; T-cell epitope; prevention; treatment;
 KW pollinosis.
 OS Chamaecyparis obtusa.
 PN J08176192-A.
 PD 09-JUL-1996.
 PF 21-DEC-1994; 335089.
 PR 21-DEC-1994; JP-335089.
 PA (MEIP) MEIJI MILK PROD CO LTD.
 DR WPI: 95-368225/37.
 CP N-BSDB: T38519.
 PT DNA encoding chamaecyparis obtusa pollen allergen: T cell
 epitope(s) of which are useful in development of preventative and
 treating agent for C. obtusa pollen pollinosis
 PS Claim 8; Pages 11-12; 17pp; Japanese.
 CC The present sequence is the C. obtusa pollen allergen Cha o 1.
 CC The T-cell epitopes of which can be used in the development of a
 CC preventive and treating agent for C. obtusa pollen pollinosis.
 CC C. obtusa pollen (2.4 kg) was degassed with diethyl ether, and
 CC dried at room temp. overnight. Cha o 1 was sepd. from it and
 CC purified. RNA was extracted from C. obtusa pollen, and mRNA and
 CC cDNA derived.
 SQ Sequence 375 AA;

Query Match 100.0%; Score 143; DB 1; Length 375;
 Best Local Similarity 100.0%; P.ed. No. 3.71e-08;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 92 IFSKNLNKLNMPLYIAGNK 111
 QY 1 IFSKNLNKLNMPLYIAGNK 20

RESULT 6
 ID R45577 standard; Protein: 367 AA.
 AC R45577.
 DI 13-JUL-1994 (first entry)
 DE Japanese cypress pollen; antigen; T-cell epitope; Chaol; Chaol2;

Japanese cedar; pollen allergen; allergy; treatment; diagnosis;
 T cell epitope; sensitivity; detection.
 OS Juniperus sabinoides.
 FH Key Location/Qualifiers
 FI peptide 1..21
 FI /note= "signal peptide"
 FI /note= "mature peptide"
 FI
 PN W09401560-A.
 PD 20-JAN-1994.
 PE 15-JAN-1993; U00139.
 PF 10-JUL-1992; W0-U05661.
 PG 01-SEP-1992; US-938990.
 PH (IMMU-) IMMUNOLOGIC PHARM CORP.
 PI Bond JF, Garman RD, Griffith IJ, Kuo M, Pollock J;
 DR WPI: 94-035065/04.
 DE N-PSDB: Q55272.
 DT Antigen(s) derived from Japanese cedar pollen allergen Cry j I -
 PI contain at least two T cell epitope(s), used to treat or diagnose
 PT allergy
 PS Disclosure: Fig 16; 137pp; English.
 CC The sequence is that of Jun s I, a homologue of the Japanese
 CC cedar pollen allergen Cry j I. Antigenic peptides derived from it
 CC can be used for the treatment and diagnosis of allergies associated
 CC with Japanese cedar pollen.
 CC
 SQ Sequence 367 AA;

Query Match 86.0%; Score 123; DB 1; Length 367;
 Best Local Similarity 75.0%; Pred. No. 8,16e-06;
 Matches 15; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

92 IFSQNNMIKLNMPLYAGNK 111
 ||| ||||| ||||| |||||
 QY 1 IFSKNLNKLNMPLYAGNK 20

RESULT 7
 ID R45578 standard; Protein: 370 AA.

AC R45578:
 DT 13-JUL-1994 (first entry)
 DE Jun v I.
 KW Japanese cedar; pollen allergen; allergy; treatment; diagnosis;
 KW T cell epitope; sensitivity; detection.
 OS Juniperus virginiana.
 FH Key Location/Qualifiers
 FI peptide 1..21
 FI /note= "signal peptide"
 FI /note= "mature peptide"
 FI
 PN W09401560-A.
 PD 20-JAN-1994.
 PE 15-JAN-1993; U00139.
 PF 10-JUL-1992; W0-U05661.
 PG 01-SEP-1992; US-938990.
 PH (IMMU-) IMMUNOLOGIC PHARM CORP.
 PI Bond JF, Garman RD, Griffith IJ, Kuo M, Pollock J;
 DR N-PSDB: Q55273.
 DT Antigen(s) derived from Japanese cedar pollen allergen Cry j I -
 PI contain at least two T cell epitope(s), used to treat or diagnose
 PT allergy
 PS Disclosure: Fig 17; 137pp; English.
 CC The sequence is that of Jun v I, a homologue of the Japanese
 CC cedar pollen allergen Cry j I. Antigenic peptides derived from it
 CC can be used for the treatment and diagnosis of allergies associated
 CC with Japanese cedar pollen.
 CC
 SQ Sequence 370 AA;

Query Match 86.0%; Score 123; DB 1; Length 370;
 Best Local Similarity 75.0%; Pred. No. 8,16e-06;
 Matches 15; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

92 IFSQNNMIKLNMPLYAGNK 111

||| ||||| ||||| |||||
 QY 1 IFSKNLNKLNMPLYAGNK 20

RESULT 8

ID R82498 standard; Protein: 20 AA.
 AC R82498:
 DT 15-APR-1996 (first entry)
 DE Cry j I Japanese cedar pollen allergen peptide fragment (CJ1-8).
 KW Cry j I; Japanese cedar pollen allergen; modified; drug production;
 KW allergy; Cryptomeria japonica.
 OS Cryptomeria japonica.
 PN W09427786-A.
 PD 19-OCT-1995.
 PE 06-APR-1995; U04249.
 PF 08-APR-1994; US-226248.
 PG 06-DEC-1994; US-350225.
 PH (IMMU-) IMMUNOLOGIC PHARM CORP.
 PI Chen X, Evans S, Franzen HM, Kuo M, Powers SP;
 PI Shaked Z;
 DR WPI: 95-36639/47.
 DT Modified Cryptomeria japonica (Cry j) I peptide(s) - useful for
 DT treating allergy to Japanese cedar pollen allergen or
 DT immunologically cross reactive allergens
 PS Disclosure: Figure 2; 60pp; English.
 CC Novel peptides of cry j I have been modified as a part of a
 CC therapeutic treatment of humans suffering from allergy to Japanese
 CC cedar pollen allergen or an allergen which is immunologically cross
 CC reactive with Japanese cedar pollen allergen. Such modified peptides
 CC possess certain characteristics which render them particularly
 CC suitable for drug product formulation. Peptide fragments of Cry j I,
 CC modified and unmodified, are given in R82491-R82525. This peptide
 CC fragment corresponds to amino acids 71-90 of the allergen mature
 CC protein.
 CC
 SQ Sequence 20 AA;

Query Match 79.0%; Score 113; DB 1; Length 20;
 Best Local Similarity 75.0%; Pred. No. 1,15e-04;
 Matches 15; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

1 IFSQNNMIKLNMPLYAGNK 111
 ||| ||||| ||||| |||||
 QY 1 IFSKNLNKLNMPLYAGNK 20

RESULT 9

ID R45549 standard; Protein: 20 AA.
 AC R45549:
 DT 13-JUL-1994 (first entry)
 DE Cry j I pollen allergen peptide CJ1-8
 KW Japanese cedar; detection; allergy; treatment; diagnosis;
 KW T cell epitope; sensitivity.
 OS Cryptomeria japonica.
 PN W09401560-A.
 PD 20-JAN-1994.
 PE 15-JAN-1993; U00139.
 PF 10-JUL-1992; W0-U05661.
 PG 01-SEP-1992; US-938990.
 PH (IMMU-) IMMUNOLOGIC PHARM CORP.
 PI Bond JF, Garman RD, Griffith IJ, Kuo M, Pollock J;
 DR WPI: 94-035066/04.
 DT Antigen(s) derived from Japanese cedar pollen allergen Cry j I -
 PI contain at least two T cell epitope(s), used to treat or diagnose
 PT allergy
 PS Claim 1; Fig 13; 137pp; English.
 CC The sequence is that of an isolated peptide of the Japanese cedar
 CC pollen allergen Cry j I (amino acids 71-90). The peptide, CJ1-8,
 CC can be used for the treatment and diagnosis of allergies associated
 CC with Japanese cedar pollen. It has enhanced therapeutic properties
 CC but reduced side effects compared to naturally occurring allergens.
 CC
 SQ Sequence 20 AA;

Query Match: 79.0%, Score 113, DB 1, Length 20;
 Best Local Similarity 75.0%, Pred. No. 1, 15e-04;
 Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

DB 1 IFSGNMNIKLKMPMYIAGYK 20
 QY 1 IFSKNLNKLKMPPLYIAGNK 20

RESULT 10
 ID W44682 standard; peptide: 30 AA.
 AC W44682;
 DT 01-MAY-1996 (first entry)
 DE T-cell epitope peptide #1 of sugi pollen antigen.
 KW T-cell epitope: sugi pollen antigen; sugi pollinosis.
 OS Synchthon;
 PS Cryptomeria japonica.
 PR J1000700-A.
 PE 13-JAN-1998.
 FR 24-JUN-1996; JP-163287.
 PA (DAIIC) DAICHI CHEM IND LTD.
 RA (MEIJI) MEIJI SEIKA KAKIASHA LTD.
 WK WPI: 98-133630/13.
 FT T cell epitope peptide of sugi pollen antigen - useful in the treatment of sugi pollinosis.
 PS Claim 1; Page 1; 14pp; Japanese.
 CC T-cell epitope peptides W44682-86 and their derivatives react with sugi pollinosis patient peripheral blood T lymphocytes. A composition prepared by combining at least 2 of the above peptides and/or their derivatives is used for the prevention and treatment of sugi pollinosis.
 CC Sequence 30 AA.

Query Match: 79.0%, Score 113, DB 1, Length 30;
 Best Local Similarity 75.0%, Pred. No. 1, 15e-04;
 Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

DB 1 IFSGNMNIKLKMPMYIAGYK 30
 QY 1 IFSKNLNKLKMPPLYIAGNK 20

RESULT 11
 ID W27371 standard; peptide: 134 AA.
 AC W27371;
 DT 24-MAR-1998 (first entry)
 DE Multi-epitope peptide used as immunotherapeutic agent #3.
 KW Multi-epitope peptide: immunotherapeutic agent; allergic diseases; T-cell epitope region; allergen; lymphocyte; immunoglobulin E.
 OS Synthetic.
 PS W09732600-A1.
 PR 12-SEP-1997.
 PE 10-MAR-1997; J00740.
 FR 10-MAR-1996; JP-080702.
 PA (MEIJI) MEIJI MILK PROD CO LTD.
 RA Dairiki K. Iwara A. Kino K. Kume A. Sone T.
 WK WPI: 97-470495/43.
 FT Peptide immunotherapeutic agent to treat allergic diseases - contains multi-epitope peptide containing T cell epitope regions from different allergens.
 PS Claim 5; Page 32; 38pp; Japanese.
 CC The present sequence represents a multi-epitope peptide which is used as a new immunotherapeutic agent. It comprises T cell epitope regions from 2 or more different allergens (preferably linked via arginine or lysine dimers), where the T cell epitope regions have a positivity index greater than 100 as measured in a patient group responding to the allergen; have at least 70% reactivity with lymphocytes from patients responding to the allergen; and are not reactive with immunoglobulin E (IgE) antibodies from patients responsive to the allergen. The agent can be used to prevent and treat a wide variety of allergic diseases, e.g. by desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
 CC Sequence 134 AA.

Query Match: 79.0%, Score 113, DB 1, Length 134;
 Best Local Similarity 75.0%, Pred. No. 1, 15e-04;
 Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

DB 78 IFSGNMNIKLKMPMYIAGYK 97
 QY 1 IFSKNLNKLKMPPLYIAGNK 20

RESULT 12
 ID R75388 standard; protein: 353 AA.
 AC R75388;
 DT 12-MAR-1996 (first entry)
 DE Japanese cedar pollen allergen
 KW Japanese cedar; pollen allergen; Cryj 1; T-cell epitope; peptides.
 OS Prevention; treatment; cryptomeria pollinosis.
 PS Cryptomeria japonica.
 PR Key Location/Qualifiers
 FT peptide 61..75
 FT /note="T-cell epitope peptide"
 FT peptide 91..105
 FT /note="T-cell epitope peptide"
 FT peptide 106..120
 FT /note="T-cell epitope peptide"
 FT peptide 146..160
 FT /note="T-cell epitope peptide"
 FT peptide 211..225
 FT /note="T-cell epitope peptide"
 FT peptide 326..340
 FT /note="T-cell epitope peptide"
 FT peptide 333..346
 FT /note="T-cell epitope peptide"
 PN J07118295-A.
 PD 09-MAY-1995.
 PR 20-OCT-1993; 262626.
 PR 20-OCT-1993; JP-262626.
 PA (MEIJI) MEIJI MILK PROD CO LTD.
 DR WPI: 95-203834/27.
 FT New cryptomeria pollen allergen T-cell epitope peptide - used for prevention, treatment and investigation of Japanese cedar pollinosis.
 PS Disclosure: Figs 1-2; App; Japanese.
 CC R75388 is the Japanese cedar pollen allergen Cryj 1, from which the T-cell epitope peptides R892..R89295 were derived. The peptides can be used for the prevention and treatment of cryptomeria pollinosis, and also for the investigation of pollinosis.
 CC Sequence 453 AA.

Query Match: 79.0%, Score 113, DB 1, Length 150;
 Best Local Similarity 75.0%, Pred. No. 1, 15e-04;
 Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

DB 71 IFSGNMNIKLKMPMYIAGYK 90
 QY 1 IFSKNLNKLKMPPLYIAGNK 20

RESULT 13
 ID R81587 standard; protein: 353 AA.
 AC R81587;
 DT 24-MAY-1996 (first entry)
 DE Cedar pollen allergen B.
 KW Cedar; pollen; allergen; immunoglobulin E; IgE; T-cell epitope;
 OS antibody; pollinosis; therapy; immunotherapy.
 PS Cryptomeria japonica.
 PR EP-700929-A2.
 PN 13-MAR-1996.
 PR 08-SEP-1995; 306295.
 PR 10-SEP-1994; JP-242137.
 PR 14-JUL-1995; JP-20021.
 PR 14-JUL-1995; JP-200204.
 PA (HAYB) HAYASHIBARA SEIBUTSU KAKAKU.
 PI Hino K. Saito S. Taniguchi Y.
 DR WPI: 96-140976/15.

PI New peptide(s) derived from cedar pollen allergens - activate
 PI allergen-specific T-cells, but not allergen-specific IgE antibodies.
 PS Claim 5: Page 41-42: 36pp: English.
 CC Synthetic peptides based on portions of cedar pollen allergens A
 CC (R81596) and B (R81587) were tested for their ability to activate
 CC cedar allergen-specific T-cells, but not allergen-specific IgE
 CC antibodies. 6 peptides (R81580-R81585) were identified as T-cell
 CC epitopes. These peptides, plus subsequences (R81573-75) essential
 CC for T-cell recognition, and homologous peptides (R81588-96) can
 CC be used as immunotherapeutic agents to treat or prevent cedar
 CC pollinosis, avoiding side-effects such as anaphylaxis.
 SQ Sequence 353 AA;

Query Match 79.0% Score 113; DB 1; Length 374;
 Best Local Similarity 75.0%; Pred. No. 115e-04;
 Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

DB 71 IFSNNNNIKKMPWYIAGYK 90
 QY 1 IFSNNNNIKKMPWYIAGNK 20

RESULT 14
 ID R81937 standard: Protein: 374 AA;
 AC R81937;

DE 03-JUN-1993 (first entry)

DE Cry j I: allergen; antigen; allergy; B cell; T cell;

DE Cryptomeria japonica;

DE Key Location/Qualifiers

DE peptide 1..21

DE /note= "s.oral peptide"

DE protein 22..374

DE /note= "mature Cry j I"

DE W09302213-A;

DE 21-JAN-1993;

DE 10-JUL-1992; 3,7661;

DE 12-JUL-1991; US-729134;

DE 15-JUL-1991; US-730452;

DE (IMMU-) IMMUNOLOGIC PHARM CORP

DE Bond JF, Griffith LJ, Pollock CJ

DE WP: 93-04544/15;

DE N-PSDB: Q35271

DE Nucleic acid sequence encoding Cryptomeria japonica allergen

DE for the diagnosis treatment and prevention of allergic reactions

DE to Japanese cedar pollen

DE Claim 11: Page 42: 69pp: English

DE Fresh pollen and staminate cone samples were collected from a single

DE Cryptomeria japonica (Japanese cedar) tree. RNA was prepd. and used

DE to synthesize cDNA. The cDNA was subjected to successive rounds of

DE PCR to yield a full length Cry j I clone. Cry j I or an antigenic

DE fragment of it may be used for detecting, treating and preventing an

DE allergic response to Japanese cedar pollen allergen. It is capable of

DE modifying both the B and T cell response to Cry j I and T cell response

DE to a Cry j I antigen.

DE Sequence 374 AA;

Query Match 79.0% Score 113; DB 1; Length 374;
 Best Local Similarity 75.0%; Pred. No. 115e-04;
 Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

DB 92 IFSNNNNIKKMPWYIAGYK 111
 QY 1 IFSNNNNIKKMPWYIAGNK 20

RESULT 15

ID R45541 standard: Protein: 374 AA;

AC R45541;

DE 23-JUL-1994 (first entry)

DE Cry j I: pollen allergen;

DE Japanese cedar; diagnosis; allergy; treatment; diagnosis;

KW I cell epitope; sensitivity;
 OS Cryptomeria japonica;
 FH Key Location/Qualifiers
 FT peptide 1..21
 FT /note= "signal peptide"
 FT peptide 22..374
 FT /note= "mature peptide"
 PN W09401560-A;
 PD 20-JAN-1994;
 PD 15-JAN-1993; U00139;
 PF 10-JUL-1992; WO-005661;
 PR 01-SEP-1992; US-938990;
 PA (IMMU-) IMMUNOLOGIC PHARM CORP
 PI Bond JF, Garman RD, Griffith LJ, Kuo M, Pollock CJ
 PI WP: 94-035056/04;
 PI N-PSDB: Q55271;
 PI Antigens derived from Japanese cedar pollen allergen Cry j I
 PI contain at least two T cell epitope(s). Used to treat or diagnose
 PI allergy
 PS Disclosure: Fig 4: 137pp: English.
 CC The sequence is that of the Japanese cedar pollen allergen
 CC Cry j I which contains at least two T cell epitopes. Peptide
 CC antigens derived from it can be used for the treatment and
 CC diagnosis of allergies associated with Japanese cedar pollen.
 CC The peptides have enhanced therapeutic properties but reduced
 CC side effects compared to naturally occurring allergens
 CC Sequence 374 AA;

Query Match 79.0% Score 113; DB 1; Length 374;
 Best Local Similarity 75.0%; Pred. No. 115e-04;
 Matches 15; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

DB 92 IFSNNNNIKKMPWYIAGYK 111
 QY 1 IFSNNNNIKKMPWYIAGNK 20

Search completed: Mon Jun 19 16:14:18 2000
 Job time : 5 secs.

□ □ □ □ □ □ □ □

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

mpsrch_pp protein - protein database search, using Smith-Waterman algorithm

```

Run on: Tue Jun 20 13:36:40 2000: MasPar time 2.38 Seconds
121.313 Million cell updates/sec
Tabular output not generated.

```

```

>US-09-142-524A-11
Description: (1-20) from US09142524A.pcp
Perfect score: 143
Sequence: : IFSKNLN:KINMPLY:AGNK 20

```

Scoring table: PAM 150
Gap 15

Searched: 145341 seqs, 14437480 residues

Post-processing: Minimum Match 08
Listing first 45

```

Database:
a-issued
1:5A_CONB 2:5B_CONB 3:6_CONB 4:PCI_CONB 5:backfiles:

```

Statistics: Mean 18.944; Variance 62.364; scale 0.396

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

S. J. M. A. B. C.

| Result No. | Score | Query | | DB | ID | Description | Prid | No. |
|------------|-------|-------|--------|------|----|------------------------|------|---------|
| | | Match | Length | | | | | |
| 1 | 64 | 44 | 6 | 383 | 1 | US-08-175-Sequence 78 | 1 | 13e-01 |
| 2 | 64 | 44 | 6 | 383 | 1 | US-08-290-Sequence 78 | 1 | 13e-01 |
| 3 | 64 | 44 | 6 | 383 | 1 | US-08-230-Sequence 78 | 1 | 13e-01 |
| 4 | 62 | 43 | 4 | 337 | 1 | US-08-175-Sequence 72 | 1 | 180e-01 |
| 5 | 62 | 43 | 4 | 387 | 1 | US-08-290-Sequence 72 | 1 | 180e-01 |
| 6 | 62 | 43 | 4 | 387 | 1 | US-08-290-Sequence 72 | 1 | 180e-01 |
| 7 | 62 | 43 | 4 | 391 | 1 | US-08-290-Sequence 59 | 1 | 180e-01 |
| 8 | 62 | 43 | 4 | 391 | 1 | US-08-290-Sequence 59 | 1 | 180e-01 |
| 9 | 62 | 43 | 4 | 391 | 1 | US-08-175-Sequence 59 | 1 | 180e-01 |
| 10 | 62 | 43 | 4 | 398 | 1 | US-08-290-Sequence 74 | 1 | 180e-01 |
| 11 | 62 | 43 | 4 | 398 | 1 | US-08-290-Sequence 74 | 1 | 180e-01 |
| 12 | 62 | 43 | 4 | 398 | 1 | US-08-290-Sequence 74 | 1 | 180e-01 |
| 13 | 59 | 41 | 3 | 397 | 1 | US-08-175-Sequence 76 | 1 | 180e-01 |
| 14 | 59 | 41 | 3 | 397 | 1 | US-08-290-Sequence 76 | 1 | 180e-01 |
| 15 | 59 | 41 | 3 | 397 | 1 | US-08-290-Sequence 76 | 1 | 180e-01 |
| 16 | 57 | 39 | 9 | 868 | 2 | US-08-783-Sequence 33 | 1 | 566e-01 |
| 17 | 57 | 39 | 9 | 868 | 2 | US-08-152-Sequence 33 | 1 | 566e-01 |
| 18 | 56 | 39 | 2 | 222 | 2 | US-05-936-Sequence 1 | 1 | 709e-01 |
| 19 | 56 | 39 | 2 | 222 | 2 | US-08-756-Sequence 1 | 1 | 709e-01 |
| 20 | 55 | 38 | 5 | 382 | 2 | US-08-728-Sequence 130 | 1 | 887e-01 |
| 21 | 55 | 38 | 5 | 382 | 2 | US-08-823-Sequence 139 | 1 | 887e-01 |
| 22 | 55 | 38 | 5 | 382 | 2 | US-08-455-Sequence 5 | 1 | 887e-01 |
| 23 | 55 | 38 | 5 | 1055 | 4 | FCI-JS96-1-Sequence 5 | 1 | 887e-01 |

CC LENGTH: 383 amino acids
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein

50 SEQUENCE 383 AA: 41852 MW: 738971 CN:

Query Match 44.8% Score 64: DB 1: Length 383;
 Best Local Similarity 35.0% Pred. No. 1.13e+01;
 Matches 7: Conservative 6: Mismatches 7: Indels 0: Gaps 0:

DB 110 IFERDMVIRLDRELAINNOK 129

QY 1 IFSKNLKNIMPLYIAGNK 20

RESULT 2

10 US-08-290-448A-78 STANDARD: PRT: 383 AA:

XX

AC xxxxxx

XX

DT

XX

DE

Sequence 78, Application US/08290448A

Sequence 78, Application US/08290448A

Patent No. 5676954

GENERAL INFORMATION:

APPLICANT: Rogers, Bruce

APPLICANT: Klappe, David G.

APPLICANT: Raina, Thorunn

APPLICANT: Kuo, Mei-chang

TITLE OF INVENTION: Allergic Proteins From Ragweed and Uses

NUMBER OF SEQUENCES: 93

CORRESPONDENCE ADDRESS:

STREET: 60 State Street, suite 510

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109-1875

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08290448A

FILING DATE: August 15, 1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/524,951

FILING DATE: May 29, 1992

APPLICATION NUMBER: US 07/325,365

FILING DATE: March 17, 1989

ATTORNEY/AGENT INFORMATION:

NAME: Atty E. Mandragoras

REGISTRATION NUMBER: 36,207

REFERENCE/DOCKET NUMBER: 2M-C18CN

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-7400

TELEFAX: (617)227-5941

INFORMATION FOR SEQ ID NO: 78:

SEQUENCE CHARACTERISTICS:

LENGTH: 383 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE 383 AA: 41852 MW: 738971 CN:

Query Match 44.8% Score 64: DB 1: Length 383;

Best Local Similarity 35.0% Pred. No. 1.13e+01;

Matches 7: Conservative 6: Mismatches 7: Indels 0: Gaps 0:

DB 110 IFERDMVIRLDRELAINNOK 129

QY 1 IFSKNLKNIMPLYIAGNK 20

RESULT 3

10 US-08-290-448A-78 STANDARD: PRT: 383 AA:

XX

AC xxxxxx

XX

DT

XX

DE

Sequence 78, Application US/08290448A

Sequence 78, Application US/08290448A

Patent No. 5698204

GENERAL INFORMATION:

APPLICANT: Rogers, Bruce

APPLICANT: Klappe, David G.

APPLICANT: Raina, Thorunn

APPLICANT: Kuo, Mei-chang

TITLE OF INVENTION: Allergic Proteins From Ragweed and Uses

NUMBER OF SEQUENCES: 93

CORRESPONDENCE ADDRESS:

STREET: 60 State Street, suite 510

CITY: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109-1875

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08290448A

FILING DATE: August 15, 1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/524,951

FILING DATE: May 29, 1992

APPLICATION NUMBER: US 07/325,365

FILING DATE: March 17, 1989

ATTORNEY/AGENT INFORMATION:

NAME: Atty E. Mandragoras

REGISTRATION NUMBER: 36,207

REFERENCE/DOCKET NUMBER: 2M-C18CN

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617)227-7400

TELEFAX: (617)227-5941

INFORMATION FOR SEQ ID NO: 78:

SEQUENCE CHARACTERISTICS:

LENGTH: 383 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE 383 AA: 41852 MW: 738971 CN:

Query Match 44.8% Score 64: DB 1: Length 383;

Best Local Similarity 35.0% Pred. No. 1.13e+01;

Matches 7: Conservative 6: Mismatches 7: Indels 0: Gaps 0:

DB 110 IFERDMVIRLDRELAINNOK 129

QY 1 IFSKNLKNIMPLYIAGNK 20

RESULT 4

10 US-08-175-069A-72 STANDARD: PRT: 387 AA:

XX

AC xxxxxx

XX

DT

XX

```

DE XX Sequence 72, Application US/08175069A
CC XX Sequence 72, Application US/08175069A
CC XX Patent No. 5776761
CC XX GENERAL INFORMATION:
CC XX APPLICANT: Rogers, Bruce
CC XX APPLICANT: Klapper, David G.
CC XX APPLICANT: Rafnar, Thorunn
CC XX APPLICANT: Kuo, Mei-chang
CC XX TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses
CC XX NUMBER OF SEQUENCES: 93
CC XX CORRESPONDENCE ADDRESS:
CC XX ADDRESSEE: LAHIVE & COCKFIELD, LLP
CC XX STREET: 60 State Street
CC XX City: Boston
CC XX STATE: Massachusetts
CC XX ZIP: 02109-1875
CC XX COMPUTER READABLE FORM:
CC XX MEDIUM TYPE: Floppy disk
CC XX OPERATING SYSTEM: PC-DOS/MS-DOS
CC XX SOFTWARE: Patent In Release #1.0, Version #1.25
CC XX CURRENT APPLICATION DATA:
CC XX FILING DATE: August 1994
CC XX PRIOR APPLICATION NUMBER:
CC XX APPLICATION NUMBER: 07/529,951
CC XX FILING DATE: May 29, 1990
CC XX APPLICATION NUMBER: US 07/325,365
CC XX FILING DATE: March 17, 1989
CC XX ATTORNEY/AGENT INFORMATION:
CC XX NAME: Amy E. Mandragouras
CC XX REGISTRATION NUMBER: 36,207
CC XX REFERENCE/DOCKET NUMBER: IMI-018CV
CC XX TELECOMMUNICATION INFORMATION:
CC XX TELEPHONE: (617)227-7400
CC XX TELEFAX: (617)227-5941
CC XX INFORMATION FOR SEQ ID NO: 72:
CC XX SEQUENCE CHARACTERISTICS:
CC XX LENGTH: 387 amino acids
CC XX TYPE: amino acid
CC XX TOPOLOGY: linear
CC XX MOLECULE TYPE: protein
CC XX SEQUENCE 387 AA: 41660 MW: 736967 CN:

Query Match 43.4% Score 62; DB 1; Length 387;
Best Local Similarity 25.0%; Pred. No. 1:80e-01;
Matches 5; Conservative 10; Mismatches 5; Indels 0; Gaps 0;

DB 109 IFEROMVIRLOKEMVNSDK 129
QY 1 IFSKNLNKLNPLYAGNK 20

RESULT 5
ID US-08-290-448A-72 STANDARD: PRT: 387 AA.
XX xxxxxx
AC
XX
DT
XX
XX
DE
XX
CC Sequence 72, Application US/08290448A
CC Patent No. 5678554
CC GENERAL INFORMATION:
CC APPLICANT: Rogers, Bruce
CC APPLICANT: Klapper, David G.
CC APPLICANT: Rafnar, Thorunn
CC APPLICANT: Kuo, Mei-chang
CC TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses
CC NUMBER OF SEQUENCES: 93

```

```

CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: LAHIVE & COCKFIELD
CC STREET: 60 State Street, suite 510
CC City: Boston
CC STATE: Massachusetts
CC COUNTRY: USA
CC ZIP: 02109-1875
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: 08/290,448A
CC FILING DATE: August 1994
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/529,951
CC FILING DATE: May 29, 1990
CC APPLICATION NUMBER: US 07/325,365
CC FILING DATE: March 17, 1989
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Amy E. Mandragouras
CC REGISTRATION NUMBER: 36,207
CC REFERENCE/DOCKET NUMBER: IMI-018CV
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617)227-7400
CC TELEFAX: (617)227-5941
CC INFORMATION FOR SEQ ID NO: 72:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 387 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 387 AA: 41660 MW: 736967 CN:

Query Match 43.4% Score 62; DB 1; Length 387;
Best Local Similarity 25.0%; Pred. No. 1:80e-01;
Matches 5; Conservative 10; Mismatches 5; Indels 0; Gaps 0;

DB 109 IFEROMVIRLOKEMVNSDK 129
QY 1 IFSKNLNKLNPLYAGNK 20

RESULT 6
ID US-08-290-448A-72 STANDARD: PRT: 387 AA.
XX xxxxxx
AC
XX
DT
XX
XX
DE
XX
CC Sequence 72, Application US/08290448A
CC Patent No. 5698204
CC GENERAL INFORMATION:
CC APPLICANT: Rogers, Bruce
CC APPLICANT: Klapper, David G.
CC APPLICANT: Rafnar, Thorunn
CC APPLICANT: Kuo, Mei-chang
CC TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses
CC NUMBER OF SEQUENCES: 93
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: LAHIVE & COCKFIELD
CC STREET: 60 State Street, suite 510
CC City: Boston
CC STATE: Massachusetts
CC COUNTRY: USA
CC ZIP: 02109-1875
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS

```


Matches 7: Conservative 7: Mismatches 6: Indels 0: Gaps 0:

00 119 11KNDVV:NLNDELWVNSDK 138

01 119 11KNDVV:NLNDELWVNSDK 138

02 119 11KNDVV:NLNDELWVNSDK 138

Search completed: Tue Jun 20 14:36:45 2000
Job time : 5 sec.

CC REFERENCE/DOCKET NUMBER: 025.6 US (IM-028CP2)
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 466-6000
CC TELEFAX: (617) 466-6040
CC INFORMATION FOR SEQ ID NO: 95:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 367 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 367 AA: 39846 MW: 695078 CN:
Query Match 86.0% Score 123: DB 10: Length 367:
Best Local Similarity 75.0% Pred. No. 1.26e-05:
Matches 15: Conservative 4: Mismatches 1: Indels 0: Caps 0:
DB 92 IFSONMNIKLMPYVAGHK 111
QY I FSKNLNLIKLNMPYVAGNK 20
RESULT 7
ID US-08-467-006-5 STANDARD PRI: 367 AA:
XX xxxxxx
AC
LI
XY
LE Sequence 95, Application: US/08457006
XX Sequence 95, Application: US/08457006
CC GENERAL INFORMATION:
CC APPLICANT: Griffith, Irwin J.
CC APPLICANT: Pollock, Joanne
CC APPLICANT: Bond, Julian F.
CC APPLICANT: Garman, Richard D.
CC APPLICANT: Kuo, Mei-Chang
CC APPLICANT: Yeung, Siu-mei H.
CC APPLICANT: Brauer, Andrew
CC APPLICANT: Exley, Mark A.
CC APPLICANT: Powers, Steven P.
CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
CC TITLE OF INVENTION: Japanese Cedar Pollen
CC NUMBER OF SEQUENCES: 261
CC CORRESPONDENCE ADDRESS:
CC ADDRESS: Immunologic Pharmaceutical Corporation, Inc.
CC STREET: 610 Lincoln St
CC CITY: Waltham
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02154
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/467,006
CC FILING DATE: June 6, 1995
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 09/350,225
CC FILING DATE: December 6, 1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Jane E. Remillard
CC REGISTRATION NUMBER: 38,872
CC REFERENCE/DOCKET NUMBER: 025.6 US06 (IM-028CP06)
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 227-7400
CC TELEFAX: (617) 227-5941
CC INFORMATION FOR SEQ ID NO: 95:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 367 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 367 AA: 39846 MW: 695078 CN:

CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 367 AA: 39846 MW: 695078 CN:
Query Match 86.0% Score 123: DB 10: Length 367:
Best Local Similarity 75.0% Pred. No. 1.26e-05:
Matches 15: Conservative 4: Mismatches 1: Indels 0: Caps 0:
DB 92 IFSONMNIKLMPYVAGHK 111
QY I FSKNLNLIKLNMPYVAGNK 20
RESULT 8
ID US-08-467-697-95 STANDARD PRI: 367 AA:
XX xxxxxx
AC
LI
XY
LE Sequence 95, Application: US/08457007
XX Sequence 95, Application: US/08457007
CC GENERAL INFORMATION:
CC APPLICANT: Griffith, Irwin J.
CC APPLICANT: Pollock, Joanne
CC APPLICANT: Bond, Julian F.
CC APPLICANT: Garman, Richard D.
CC APPLICANT: Kuo, Mei-Chang
CC APPLICANT: Yeung, Siu-mei H.
CC APPLICANT: Brauer, Andrew
CC APPLICANT: Exley, Mark A.
CC APPLICANT: Powers, Steven P.
CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
CC TITLE OF INVENTION: Japanese Cedar Pollen
CC NUMBER OF SEQUENCES: 261
CC CORRESPONDENCE ADDRESS:
CC ADDRESS: Immunologic Pharmaceutical Corporation, Inc.
CC STREET: 610 Lincoln St
CC CITY: Waltham
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02154
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/467,697
CC FILING DATE: June 6, 1995
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/350,425
CC FILING DATE: December 6, 1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Jane E. Remillard
CC REGISTRATION NUMBER: 38,872
CC REFERENCE/DOCKET NUMBER: 025.6 US04 (IM-028CP04)
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 227-7400
CC TELEFAX: (617) 227-5941
CC INFORMATION FOR SEQ ID NO: 95:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 367 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 367 AA: 39846 MW: 695078 CN:
Query Match 86.0% Score 123: DB 10: Length 367:
Best Local Similarity 75.0% Pred. No. 1.26e-05:

Matches 15: Conservative 4: Mismatches 1: Indels 0: Gaps 0:

DE 92 IFSQNNIKLNKMPYVAGHK 111
 QY 1 IFSKNNIKLNKMPYIAGNK 20

RESULT 9
 ID US-08-467-021-47 STANDARD: PRI: 370 AA.
 XX
 AC xxxxxx

Sequence 97, Application US/08467023

Sequence 97, Application US/08467023

GENERAL INFORMATION:
 APPLICANT: Griffith, Irwin J.
 APPLICANT: Pollock, Joanne
 APPLICANT: Bond, Julian F.
 APPLICANT: Garmar, Richard D.
 APPLICANT: Kuo, Mei-Chang
 APPLICANT: Yeung, Siu-mei H.
 APPLICANT: Brauer, Andrew
 APPLICANT: Exley, Mark A.
 APPLICANT: Powers, Steven P.
 TITLE OF INVENTION: Allergenic Proteins And Peptides From
 TITLE OF INVENTION: Japanese Cedar Pollen
 NUMBER OF SEQUENCES: 261
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: ImmuLogic Pharmaceutical Corporation, Inc.
 STREET: 610 Lincoln St
 CITY: Waltham
 STATE: MA
 COUNTRY: USA
 ZIP: 02154

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/467,023
 FILING DATE: June 6, 1995
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/350,225
 FILING DATE: December 6, 1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Jane E. Rehill and
 REGISTRATION NUMBER: 38,872
 REFERENCE/DOCKET NUMBER: 025.6 USD2 (IMI-028CPD2)

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 227-7400
 TELEFAX: (617) 227-5941
 INFORMATION FOR SEQ ID NO: 97:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 370 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE 370 AA: 40191 MW: 725506 CN:

Query Match 86.0% Score 123; DB 10; Length 370;
 Best Local Similarity 75.0% Pred. No. 1.26e-05;
 Matches 15; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

DE 92 IFSQNNIKLNKMPYVAGHK 111
 QY 1 IFSKNNIKLNKMPYIAGNK 20

RESULT 10
 ID US-08-467-697-97 STANDARD: PRI: 370 AA.
 XX
 AC xxxxxx

Sequence 97, Application US/08467697

Sequence 97, Application US/08467697

GENERAL INFORMATION:
 APPLICANT: Griffith, Irwin J.
 APPLICANT: Pollock, Joanne
 APPLICANT: Bond, Julian F.
 APPLICANT: Garmar, Richard D.
 APPLICANT: Kuo, Mei-Chang
 APPLICANT: Yeung, Siu-mei H.
 APPLICANT: Brauer, Andrew
 APPLICANT: Exley, Mark A.
 APPLICANT: Powers, Steven P.
 TITLE OF INVENTION: Allergenic Proteins And Peptides From
 TITLE OF INVENTION: Japanese Cedar Pollen
 NUMBER OF SEQUENCES: 261
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: ImmuLogic Pharmaceutical Corporation, Inc.
 STREET: 610 Lincoln St
 CITY: Waltham
 STATE: MA
 COUNTRY: USA
 ZIP: 02154

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/467,697
 FILING DATE: June 6, 1995
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 350,225
 FILING DATE: December 6, 1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Jane E. Rehill and
 REGISTRATION NUMBER: 38,872
 REFERENCE/DOCKET NUMBER: 025.6 USD4 (IMI-028704)

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 227-7400
 TELEFAX: (617) 227-5941
 INFORMATION FOR SEQ ID NO: 97:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 370 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE 370 AA: 40191 MW: 725506 CN:

Query Match 86.0% Score 123; DB 10; Length 370;
 Best Local Similarity 75.0% Pred. No. 1.26e-05;
 Matches 15; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

DE 92 IFSQNNIKLNKMPYVAGHK 111
 QY 1 IFSKNNIKLNKMPYIAGNK 20

RESULT 11
 ID US-08-226-249A-97 STANDARD: PRI: 370 AA.
 XX
 AC xxxxxx

```

DE Sequence 97, Application US/08226248A
XX
CC Sequence 97, Application US/08226248A
CC GENERAL INFORMATION:
CC APPLICANT: Griffith, Irwin J.
CC APPLICANT: Pollock, Joanne
CC APPLICANT: Bond, Julian
CC APPLICANT: Garman, Richard D.
CC APPLICANT: Kuo, Mei-Chang
CC APPLICANT: Yeung, Siu-mei H.
CC APPLICANT: Brauer, Andrew
CC APPLICANT: Exley, Mark A.
CC APPLICANT: Powers, Steven P.
CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
CC TITLE OF INVENTION: Japanese Cedar Pollen
CC NUMBER OF SEQUENCES: 26
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
CC STREET: 610 Lincoln St.
CC CITY: Waltham
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02154
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/226,248A
CC FILING DATE: April 8, 1994
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/350,225
CC FILING DATE: December 6, 1994
CC APPLICATION NUMBER: 08/226,248
CC FILING DATE: April 8, 1994
CC APPLICATION NUMBER: 07/938,990
CC FILING DATE: September 1, 1992
CC APPLICATION NUMBER: PCT/US93/00139
CC FILING DATE: January 15, 1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Darlene A. Vanstone
CC REGISTRATION NUMBER: 35,729
CC REFERENCE/DOCKET NUMBER: 025.5 US (IM1-028CP)
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 466-6000
CC TELEFAX: (617) 466-6040
CC INFORMATION FOR SEQ ID NO: 97:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 370 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE: 370 AA; 4319 MW; 725506 CN;

Query Match 86.0%; Score 123; DB 7; Length 370;
Best Local Similarity 75.0%; Pred. No. 1,26e-05;
Matches 15; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 92 IFSONNKKLNKMPLYVAGHK 111
QY 1 IFSKNKLNKLNKMPLYIAGNK 20

RESULT 12
ID US-09-468-940-97 STANDARD: PRT: 370 AA.
XX
XX xxxxxx
XX
XX
XX
XX
DE Sequence 97, Application US/08468940
XX
CC Sequence 97, Application US/08468940
CC GENERAL INFORMATION:
CC APPLICANT: Griffith, Irwin J.
CC APPLICANT: Pollock, Joanne
CC APPLICANT: Bond, Julian
CC APPLICANT: Garman, Richard D.
CC APPLICANT: Kuo, Mei-Chang
CC APPLICANT: Yeung, Siu-mei H.
CC APPLICANT: Brauer, Andrew
CC APPLICANT: Exley, Mark A.
CC APPLICANT: Powers, Steven P.
CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
CC TITLE OF INVENTION: Japanese Cedar Pollen
CC NUMBER OF SEQUENCES: 26
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
CC STREET: 610 Lincoln St.
CC CITY: Waltham
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02154
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/226,248A
CC FILING DATE: April 8, 1994
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/938,990
CC FILING DATE: September 1, 1992
CC APPLICATION NUMBER: PCT/US93/00139
CC FILING DATE: January 15, 1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Darlene A. Vanstone
CC REGISTRATION NUMBER: 35,729
CC REFERENCE/DOCKET NUMBER: 025.5 US (IM1-028CP)
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 466-6000
CC TELEFAX: (617) 466-6040
CC INFORMATION FOR SEQ ID NO: 97:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 370 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE: 370 AA; 4319 MW; 725506 CN;

Query Match 86.0%; Score 123; DB 7; Length 370;
Best Local Similarity 75.0%; Pred. No. 1,26e-05;
Matches 15; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 92 IFSONNKKLNKMPLYVAGHK 111
QY 1 IFSKNKLNKLNKMPLYIAGNK 20

RESULT 13
ID US-09-467-006-97 STANDARD: PRT: 370 AA.
XX
XX xxxxxx
XX
XX
XX
XX
DE Sequence 97, Application US/08467006
XX
CC Sequence 97, Application US/08467006
CC GENERAL INFORMATION:
CC APPLICANT: Griffith, Irwin J.
CC APPLICANT: Pollock, Joanne;

```


CC APPLICANT: Bond, Julian F.;
 CC APPLICANT: Garman, Richard D;
 CC APPLICANT: Kuo, Mei-Chang; H.;
 CC APPLICANT: Yeung, Siu-mei H.;
 CC APPLICANT: Brauer, Andrew;
 CC APPLICANT: Exley, Mark A.;
 CC APPLICANT: Powers, Steven P.;
 CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
 CC TITLE OF INVENTION: Japanese Cedar Pollen
 CC NUMBER OF SEQUENCES: 261
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
 CC STREET: 610 Lincoln St
 CC CITY: Waltham
 CC STATE: MA
 CC COUNTRY: USA
 CC ZIP: 02154
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: PatentIn Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/467,006
 CC FILING DATE: June 6, 1995
 CC CLASSIFICATION: 424
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/350,225
 CC FILING DATE: December 6, 1994
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Jane E. Remillard
 CC REGISTRATION NUMBER: 36,872
 CC REFERENCE/DOCKET NUMBER: 025.5 US/09 (MI-028CPD6)
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (617) 227-7400
 CC TELEFAX: (617) 227-5941
 CC INFORMATION FOR SEQ ID NO: 97:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 370 amino acids
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC SEQUENCE 370 AA: 40191 MW: 725506 CN:

Query Match 86.0% Score 123.08 bits Length 370
 Best Local Similarity 75.0% Prod No: 126-05
 Matches 15: Conservative 4: Missed 4: Indels 0: Gaps 0:

CU 92 IFSNNKIKKMPLYVAGHK 111
 CY 1 IFSKNLNKLNMPLYAGNK 20

RESULT 14
 ID US-08-350-225-97 STANDARD: PRT: 370 AA.

XX xxxxxx

Sequence 97. Application US/08350225

Sequence 97. Application US/08350225

GENERAL INFORMATION:
 CC APPLICANT: Griffith, Irwin J.;
 CC APPLICANT: Pollock, Joanne;
 CC APPLICANT: Bond, Julian F.;
 CC APPLICANT: Garman, Richard D;
 CC APPLICANT: Kuo, Mei-Chang;
 CC APPLICANT: Yeung, Siu-mei H.;
 CC APPLICANT: Brauer, Andrew;
 CC APPLICANT: Exley, Mark A.;
 CC APPLICANT: Powers, Steven P.

CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
 CC TITLE OF INVENTION: Japanese Cedar Pollen
 CC NUMBER OF SEQUENCES: 261
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
 CC STREET: 610 Lincoln St
 CC CITY: Waltham
 CC STATE: MA
 CC COUNTRY: USA
 CC ZIP: 02154
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: PatentIn Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/350,225
 CC FILING DATE: December 6, 1994
 CC CLASSIFICATION: 424
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/226,248
 CC FILING DATE: April 8, 1994
 CC APPLICATION NUMBER: 07/938,990
 CC FILING DATE: September 1, 1992
 CC APPLICATION NUMBER: PCT/US93/00149
 CC FILING DATE: January 15, 1993
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Darlene A. Vanstone
 CC REGISTRATION NUMBER: 35,729
 CC REFERENCE/DOCKET NUMBER: 025.6 US (MI-028CP2)
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (617) 466-6030
 CC TELEFAX: (617) 466-6040
 CC INFORMATION FOR SEQ ID NO: 97:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 370 amino acids
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC SEQUENCE 370 AA: 40191 MW: 745506 CN:

Query Match 86.0% Score 123.08 bits Length 370
 Best Local Similarity 75.0% Prod No: 126-05
 Matches 15: Conservative 4: Missed 4: Indels 0: Gaps 0:

CU 92 IFSNNKIKKMPLYVAGHK 111
 CY 1 IFSKNLNKLNMPLYAGNK 20

RESULT 15
 ID US-08-467-006-33 STANDARD: PRT: 20 AA.

XX xxxxxx

Sequence 33. Application US/08467006

Sequence 33. Application US/08467006

GENERAL INFORMATION:
 CC APPLICANT: Griffith, Irwin J.;
 CC APPLICANT: Pollock, Joanne;
 CC APPLICANT: Bond, Julian F.;
 CC APPLICANT: Garman, Richard D;
 CC APPLICANT: Kuo, Mei-Chang;
 CC APPLICANT: Yeung, Siu-mei H.;
 CC APPLICANT: Brauer, Andrew;
 CC APPLICANT: Exley, Mark A.;
 CC APPLICANT: Powers, Steven P.
 CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
 CC TITLE OF INVENTION: Japanese Cedar Pollen
 CC NUMBER OF SEQUENCES: 261

CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Immunologic Pharmaceuticals Corporation, Inc.
 CC STREET: 610 Lincoln St
 CC CITY: Waltham
 CC STATE: MA
 CC COUNTRY: USA
 CC ZIP: 02154
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patent Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/467,006
 CC FILING DATE: June 6, 1995
 CC CLASSIFICATION: 424
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/350,225
 CC FILING DATE: December 6, 1994
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Jane E. Remillard
 CC REGISTRATION NUMBER: 38,872
 CC REFERENCE/DOCKET NUMBER: 025.6 USD6 (IM-028CP06)
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (617) 227-7400
 CC TELEFAX: (617) 227-5941
 CC INFORMATION FOR SEQ ID NO: 33:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 20 amino acids
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: peptide
 CC FRAGMENT TYPE: internal
 CC SEQUENCE 20 AA: 2320 MW: 2402 CN:

Query Match 79.0% Score 113 DB 10 Length 20
 Best Local Similarity 75.0% Pref. No. 2,25e-04
 Matches 15: Conservative 3: Mismatches 2: Indels 0: Gaps 0:

15 1 TFSNNRIKLNMPYIAGNK 20
 17 1 TFSNNRIKLNMPYIAGNK 20

Search completed: Mon Jun 19 16:27:27 2000
 Ser time : 19 secs.

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

```

search_pprotein - protein database search, using Smith-Waterman algorithm
run on: M05 Jun 19 16:13:48 2000: MasPar time 5.53 Seconds
        170 634 Million cell updates/sec
bular output not generated.

```

```
>US-09-142-524A-11  
Description: (1-20) from US09142524A.rep  
Effect Score:  
Success: 1 FSKYIN:KLNWP-Y:AGNK 20
```

oring table: PAM :50
Gap :5

searched: 142060 seqs, 47172406 residues

st-processing: Minimum Match 08
listing first 45 summaries

```
base: 01r62
```

artist:os: Mean 29.146; Variance 45.738; scale 0.637

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARY

| Alt | No. | Score | Query Match | Length | EF | ID | Description | Pred. No. |
|-----|-----|-------|-------------|--------|--------|--------------------------------|-------------|-----------|
| 1 | 113 | 79.0 | 374 | 2 | 7C2124 | major allergen Cry j | 2.35e-09 | |
| 2 | 113 | 79.0 | 374 | 2 | 7C2123 | major allergen Cry j | 2.35e-09 | |
| 3 | 81 | 56.6 | 347 | 2 | S62611 | pectate lyase (EC 4.2.1.1) | 5.52e-03 | |
| 4 | 74 | 51.7 | 347 | 2 | S59837 | probable membrane protein | 1.02e-01 | |
| 5 | 67 | 47.6 | 455 | 2 | T00366 | pectate lyase (EC 4.2.1.1) | 1.11e-00 | |
| 6 | 57 | 46.3 | 392 | 2 | D53420 | allergen Amb a 1.4 | 1.63e-00 | |
| 7 | 56 | 46.2 | 434 | 2 | S29612 | pectate lyase (EC 4.2.1.1) | 2.40e-00 | |
| 8 | 63 | 44.1 | 350 | 2 | E64320 | hypothetical protein | 5.09e-00 | |
| 9 | 63 | 44.1 | 150 | 2 | E64638 | conserved hypothetical protein | 7.38e-00 | |
| 10 | 63 | 44.1 | 150 | 2 | D71575 | hypothetical protein | 7.38e-00 | |
| 11 | 63 | 44.1 | 452 | 2 | D47236 | ribulose-bisphosphate | 7.38e-00 | |
| 12 | 52 | 43.4 | 155 | 1 | B64250 | osmotically inducible | 1.06e-01 | |
| 13 | 62 | 43.4 | 201 | 1 | DAP3AA | proteoglycan 3,4-d | 1.06e-01 | |
| 14 | 62 | 43.4 | 396 | 2 | A39099 | allergen Amb a 1.1 pr | 1.06e-01 | |
| 15 | 62 | 43.4 | 398 | 2 | B39099 | allergen Amb a 1.2 | 1.06e-01 | |
| 16 | 62 | 43.4 | 398 | 2 | B53240 | allergen Amb a 1.2 pr | 1.06e-01 | |
| 17 | 62 | 43.4 | 398 | 2 | T07058 | pectate lyase (EC 4.2.1.1) | 1.06e-01 | |
| 18 | 62 | 43.4 | 618 | 2 | G70350 | conserved hypothetical protein | 1.06e-01 | |
| 19 | 62 | 43.4 | 1956 | 2 | T16416 | hypothetical protein | 1.06e-01 | |
| 20 | 61 | 42.7 | 115 | 2 | U58950 | NADH dehydrogenase (u | 1.53e-01 | |
| 21 | 61 | 42.7 | 115 | 2 | A41827 | NADH dehydrogenase (u | 1.53e-01 | |
| 22 | 61 | 42.7 | 131 | 2 | G71243 | hypothetical protein | 1.53e-01 | |
| 23 | 61 | 42.7 | 172 | 2 | PC4440 | aluminum tolerance pr | 1.53e-01 | |

| | | | | | | | |
|----|----|------|------|---|--------|--------------------------|---------|
| 24 | 61 | 42.7 | 236 | 2 | T03134 | GTP-binding protein h | 1.5e-01 |
| 25 | 61 | 42.7 | 353 | 2 | D5001 | conserved hypothetical | 1.5e-01 |
| 26 | 61 | 42.7 | 649 | 2 | S5047 | deoxyribosyl-transferase | 1.5e-01 |
| 27 | 61 | 42.7 | 1337 | 2 | B6993 | hypothetical protein | 1.5e-01 |
| 28 | 60 | 42.0 | 214 | 2 | T25692 | hypothetical protein | 2.1e-01 |
| 29 | 60 | 42.0 | 274 | 2 | T03932 | pectate lyase (EC 4.2 | 2.1e-01 |
| 30 | 60 | 42.0 | 275 | 2 | T27012 | pectate lyase - pseudo | 2.1e-01 |
| 31 | 61 | 42.0 | 385 | 2 | T27459 | nirX protein - EC4.2 | 2.1e-01 |
| 32 | 60 | 42.0 | 449 | 2 | T27459 | probable type II DNA | 2.1e-01 |
| 33 | 60 | 42.0 | 454 | 2 | S21998 | pectate lyase (EC 4.2 | 2.1e-01 |
| 34 | 60 | 42.0 | 454 | 2 | A41803 | ribulose-bisphosphate | 2.1e-01 |
| 35 | 60 | 42.0 | 459 | 2 | T01640 | ribulose-bisphosphate | 2.1e-01 |
| 36 | 60 | 42.0 | 475 | 1 | RKXHLG | ribulose-bisphosphate | 2.1e-01 |
| 37 | 60 | 42.0 | 475 | 1 | RK32LK | ribulose-bisphosphate | 2.1e-01 |
| 38 | 60 | 42.0 | 475 | 1 | RK5ZLM | ribulose-bisphosphate | 2.1e-01 |
| 39 | 60 | 42.0 | 475 | 1 | RKXHLI | ribulose-bisphosphate | 2.1e-01 |
| 40 | 60 | 42.0 | 475 | 2 | A4161 | ribulose-bisphosphate | 2.1e-01 |
| 41 | 60 | 42.0 | 475 | 2 | A41615 | ribulose-bisphosphate | 2.1e-01 |
| 42 | 50 | 42.0 | 478 | 1 | RKNULI | ribulose-bisphosphate | 2.1e-01 |
| 43 | 60 | 42.0 | 480 | 1 | RK1ELP | ribulose-bisphosphate | 2.1e-01 |
| 44 | 60 | 42.0 | 523 | 2 | S61743 | carboxypeptidase Y pr | 2.1e-01 |
| 45 | 59 | 41.3 | 1478 | 2 | S74131 | DNA-directed RNA pol | 3.1e-01 |

APPENDICES

```

RESULT 1
ENTRY
  TITLE
    cedar
  ORGANISM
    Cryptomeria japonica
  DATE
    14-Jul-1994
  #sequence_revision 14-Jul-1994
  #text_change 26-Aug-1999
  JC2124
  #type complete
  #allergen Cry j 1 precursor (clone pCry-15) - Japanese cedar
  #formal_name Cryptomeria japonica
  #common_name Japanese Cedar
  #accession
    JC2124
    JC2123
  #authors
    Sone, T.; Komiyama, N.; Shimizu, K.; Kusakato, T.; Morikubo, K.; Kino, K
  #journal
    Biochem. Biophys. Res. Commun. (1994) 199:619-625
  #cite
    Cloning and sequencing of cDNA coding for Cry j 1, a major allergen of Japanese cedar pollen.
  #cross-references
    MIM:941833
  #accession
    JC2124
  #molecule_type mRNA
  #residues 1-374
  #cross-references
    G8:026475; NID:0493633; PMID:01006087; PMID:0493634
  #experimental_source pollen
  #note
    The authors described an alternative building site for

```

```

CLASSIFICATION
KEYWORDS
FEATURE
1-21
22-374
158,191,293,354
SUMMARY
length 374 #molecular-weight 40702 #checksum 1692

Query Match 79.0% Score 113; DB 2; Length 374;
Best Local Similarity 75.0%; P-val. No. 2.35e-09;
Matches 15; Conservative Mismatches 2; Indels 0; Gaps 0;

```

```

dbb 92 IFSGNMNKLMPWYIAGYK 111
    III I:IIII:I:III I
QY 1 IFSKNLNKLMPLYIAGNK 20

```

```
RESULT      2
ENTRY       JC2123      #type complete
TITLE       major allergen Cry j I precursor (clone pCCi-2-2) - Japanese cedar
DATE        #formal_name Cryptomeria japonica #common_name Japanese cedar
ORIGIN      14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change
```



```

CLASSIFICATION  *superfamily pectate lyase LATS9
KEYWORDS         carbon-oxygen lyase
SUMMARY          *length 455 *molecular-weight 5257 *checksum 611

Query Match      47.6% Score 68: DB 2: Length 455:
Best Local Similarity 40.0%: Pred. No. 1110-00:
Matches          8: Conservative 6: Mismatches 5: Indels 0: Gaps 0:

DB 171 IFARSMVIRKQELIINDK 190
||||| ||| |||
QY 1 IFSKNLIKLNMPLYAGNK 20

RESULT 6
ENTRY   *type complete
TITLE   *allogene Amb a 1.4 precursor - common ragweed
ORGANISM *normal_name Ambrosia artemisiifolia *common_name common ragweed
DATE    19-May-1994 *sequence_revision 19-May-1994 *text_change 20
REFERENCE
AUTHORS Griffith, I.J.: Pollock, J.: Klapper, D.G.: Rogers, R.L.: Nault, A.K.
JOURNAL Int. Arch. Allergy Appl. Immunol. (1991) 96:296-304
TITLE    Sequence polymorphism of Amb a 1 and Amb a II, the major allergens in Ambrosia artemisiifolia (short ragweed).
CROSS-REFERENCES MUID:92234570
ACCESSION D33240
STATUS    preliminary
MOLECULE_TYPE mRNA
RESIDUES 1-392 **label GRI
CROSS-REFERENCES GB:M80562: NID:g156444: PIDN:AAA32670.1: PID:g16645

CLASSIFICATION  *superfamily pectate lyase LATS9
KEYWORDS         pollen
SUMMARY          *length 392 *molecular-weight 42842 *checksum 9019

Query Match      45.9% Score 67: DB 2: Length 392:
Best Local Similarity 35.0%: Pred. No. 1430-00:
Matches          7: Conservative 7: Mismatches 5: Indels 0: Gaps 0:

DB 119 IFARSMVIRKQELIINDK 138
||||| ||| |||
QY 1 IFSKNLIKLNMPLYAGNK 20

RESULT 7
ENTRY   *type complete
TITLE   Pectate lyase (EC 4.2.2.2) - trumpet lily
ORGANISM *normal_name Lilium longiflorum *common_name trumpet lily
DATE    29-Mar-1997 *sequence_revision 19-Mar-1997 *text_change 20
REFERENCE
AUTHORS Kim, S.: Finkel, D.J.: An, G.
SUBMISSION Submitted to the EMBL Data Library, October 1992
DESCRIPTION Abundancy patterns of lily pollen cDNAs: characterization of three pollen-preferential cDNA clones.
ACCESSION S29612
REFERENCE
AUTHORS Kim, S.: Finkel, D.J.: An, G.
SUBMISSION Submitted to the EMBL Data Library, October 1992
DESCRIPTION Abundancy patterns of lily pollen cDNAs: characterization of three pollen-preferential cDNA clones.
ACCESSION S29612
STATUS    preliminary
MOLECULE_TYPE mRNA
RESIDUES 1-434 **label KIM
CROSS-REFERENCES EMBL:Z17328: NID:g19450: PIDN:CAAT8976.1: PID:g19451
EXPERIMENTAL_SOURCE cv. Nellie White, mature flower
CLASSIFICATION  *superfamily pectate lyase LATS9
KEYWORDS         carbon-oxygen lyase
SUMMARY          *length 434 *molecular-weight 48457 *checksum 6432

Query Match      45.2% Score 66: DB 2: Length 434:
Best Local Similarity 40.0%: Pred. No. 2400-00:
Matches          8: Conservative 6: Mismatches 6: Indels 0: Gaps 0:

DB 156 IFKSMVIRKQELIINDK 175

```

```

##molecule_type DNA
##residues 1-150 ##label TOM
##cross-references GB:AE00604; GB:AE00604; NID:g2314086; PID:g2314086;
TIGR:R0949
CLASSIFICATION ##superfamily conserved hypothetical protein H10033
SUMMARY ##length 150 ##molecular-weight 17331 ##checksum 8767

Query Match 44.1% Score 63; DB 2; Length 150;
Best Local Similarity 52.9%; Pred. No. 7,350-00;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

DB 85 FSKMLENHNINFFIAG 101
QY 2 FSKNKLKLNKMPLYIAG 18

RESULT 11
ENTRY
TITLE
ORGANISM
VARIETY
DATE
ACCESSIONS
REFERENCE
AUTHORS
DESCRIPTION
JOURNAL
TITLE
CROSS-REFERENCES
ACCESSION
STATUS
MOLECULE_TYPE DNA
RESIDUES 1-150 ##label ARN
CROSS-REFERENCES GB:AE00518; GB:AE00439; NID:g4155454; PID:g4155454
EXPERIMENTAL_SOURCE strain 399
GENETICS
GENE
SUMMARY

Query Match 44.1% Score 63; DB 2; Length 150;
Best Local Similarity 52.9%; Pred. No. 7,350-00;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

DB 85 FSKMLENHNINFFIAG 101
QY 2 FSKNKLKLNKMPLYIAG 18

RESULT 11
ENTRY
TITLE
ORGANISM
DATE
ACCESSIONS
REFERENCE
AUTHORS
SUBMISSION
DESCRIPTION
ACCESSION
MOLECULE_TYPE DNA
RESIDUES 1-452 ##label MAN2
CROSS-REFERENCES EMBL:X81104

```

```

REFERENCE
S39270
Manen, J.F.; Natali, A.; Ehrendorfer, F.
Submitted to the EMBL Data Library, December 1994
Phylogeny of Rubiaceae-Rubia inferred from the sequence of a
cpDNA intergenic region.
S39292
##molecule_type DNA
##residues 1-52, X' 54 57 ##label MAN2
##cross-references EMBL:X73; S; NID:g44895; PID:g44895;
PID:g44895
GENETICS
GENE
GENOME
FUNCTION
DESCRIPTION
CLASSIFICATION
KEYWORDS
FEATURE
175-334
201
203
SUMMARY
Query Match 44.1% Score 64; DB 2; Length 452;
Best Local Similarity 33.3%; Pred. No. 7,350-00;
Matches 6; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
QY 1 FSKNKLKLNKMPLYIAG 18
DB 255 VFARLAPVIVWHDVIG 272
QY 1 FSKNKLKLNKMPLYIAG 18

RESULT 12
ENTRY
TITLE
ORGANISM
DATE
ACCESSIONS
REFERENCE
AUTHORS
JOURNAL
TITLE
CROSS-REFERENCES
ACCESSION
STATUS
MOLECULE_TYPE DNA
RESIDUES 1-155 ##label IIGR
CROSS-REFERENCES GB:U39732; GB:I43967; NID:g1046169; PID:g1046169;
TIGR:MG454
EXPERIMENTAL_SOURCE strain G-37
GENETICS
GENE
SUMMARY
Query Match 43.4% Score 62; DB 1; Length 155;
Best Local Similarity 57.1%; Pred. No. 1,06e+01;

```

 W A T E R M A N
 (TM)

Release 3.1A John F. Collins, BioComputing Research Unit.
 Copyright (C) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Jun 19 16:12:55 2000; MasPar time 3.84 seconds
 158,450 Million cell updates/sec
 Tabular output not generated.

Title: >US-09-142-524A-11
 Description: (1-20); from US09142524A.pet
 Perfect Score: 143
 Sequence: 1:FSKNLN:KLNMPPLY:AGNK 20

Scoring table: PAM 150
 Gap 15

Searched: 81257 seqs, 30454973 residues

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database: swiss-prot38
 1:swissprot

Statistics: Mean 29.800; Variance 43.129; scale 0.691

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed.
 and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description | Pred. No. |
|------------|-------|-------------|--------|----|------------------------------------|-----------|
| 1 | 123 | 79.0 | 374 | 1 | SBP_CRYJA SUGI BASIC PROTEIN PRE | 2,410-10 |
| 2 | 81 | 55.6 | 327 | 1 | PEL_TORAC PECCATE LYASE PRECURSO | 1,558-03 |
| 3 | 74 | 51.7 | 347 | 1 | RH2_YEAST DNA DAMAGE TOLERANCE P | 2,498-02 |
| 4 | 67 | 46.9 | 392 | 1 | MP4_AMBAR POLLEN ALLERGEN AMB A | 5,668-01 |
| 5 | 66 | 46.2 | 434 | 1 | PEL_LILLO PECCATE LYASE PRECURSO | 1,008-00 |
| 6 | 54 | 44.8 | 395 | 1 | Y184_METCA HYPOTHETICAL PROTEIN M | 2,278-00 |
| 7 | 63 | 44.1 | 453 | 1 | RBL_RUBP1 RIBULOSE BISPHOSPHATE | 3,318-00 |
| 8 | 63 | 44.1 | 455 | 1 | RBL_LUPM1 RIBULOSE BISPHOSPHATE | 3,318-00 |
| 9 | 63 | 44.1 | 475 | 1 | RBL_CERGL RIBULOSE BISPHOSPHATE | 3,318-00 |
| 10 | 63 | 44.1 | 475 | 1 | RBL_STEME RIBULOSE BISPHOSPHATE | 3,318-00 |
| 11 | 63 | 44.1 | 476 | 1 | RBL_AREDR RIBULOSE BISPHOSPHATE | 3,318-00 |
| 12 | 63 | 44.1 | 477 | 1 | RBL_AVEDA RIBULOSE BISPHOSPHATE | 3,318-00 |
| 13 | 62 | 43.4 | 155 | 1 | Y454_MYCSE HYPOTHETICAL PROTEIN M | 4,898-00 |
| 14 | 62 | 43.4 | 200 | 1 | PCXA_PSEPU PROTOCATECHUATE 3,4-DI | 4,898-00 |
| 15 | 62 | 43.4 | 389 | 1 | TP6A_SULSH TYPE II DNA TOPOISOMER | 4,898-00 |
| 16 | 62 | 43.4 | 396 | 1 | MP11_AMBAR POLLEN ALLERGEN AMB A | 4,898-00 |
| 17 | 62 | 43.4 | 398 | 1 | PE55_LYCES PROBABLE PECCATE LYASE | 4,898-00 |
| 18 | 62 | 43.4 | 398 | 1 | MP12_AMBAR POLLEN ALLERGEN AMB A | 4,898-00 |
| 19 | 62 | 43.4 | 1687 | 1 | VIT2_FUNHE VITELLOGENIN II PRECUR | 4,898-00 |
| 20 | 51 | 42.7 | 115 | 1 | N3M_BALPH NADH:UBIQUINONE OXIDOR | 7,178-00 |
| 21 | 51 | 42.7 | 353 | 1 | TP6A_METTH TYPE II: DNA TOPOISOMER | 7,178-00 |
| 22 | 51 | 42.7 | 469 | 1 | RBL_HALRB RIBULOSE BISPHOSPHATE | 7,178-00 |
| 23 | 60 | 42.0 | 253 | 1 | RBL_WAGLA RIBULOSE BISPHOSPHATE | 1,058-01 |

| | | | | | | |
|----|----|------|-----|---|-----------------------------------|----------|
| 24 | 60 | 42.0 | 441 | 1 | BL_CORSA RIBULOSE BISPHOSPHATE | 1,058-01 |
| 25 | 60 | 42.0 | 443 | 1 | RBL_IRIGE RIBULOSE BISPHOSPHATE | 1,058-01 |
| 26 | 60 | 42.0 | 443 | 1 | RBL_VERBC RIBULOSE BISPHOSPHATE | 1,058-01 |
| 27 | 60 | 42.0 | 443 | 1 | RBL_AB:FI RIBULOSE BISPHOSPHATE | 1,058-01 |
| 28 | 60 | 42.0 | 443 | 1 | RBL_RIVE RIBULOSE BISPHOSPHATE | 1,058-01 |
| 29 | 60 | 42.0 | 443 | 1 | RBL_ABISA RIBULOSE BISPHOSPHATE | 1,058-01 |
| 30 | 60 | 42.0 | 443 | 1 | RBL_ASHO RIBULOSE BISPHOSPHATE | 1,058-01 |
| 31 | 60 | 42.0 | 443 | 1 | RBL_CALDI RIBULOSE BISPHOSPHATE | 1,058-01 |
| 32 | 60 | 42.0 | 449 | 1 | PE59_LYCES PROBABLE PECCATE LYASE | 1,058-01 |
| 33 | 60 | 42.0 | 450 | 1 | RBL_CRAMA RIBULOSE BISPHOSPHATE | 1,058-01 |
| 34 | 60 | 42.0 | 459 | 1 | RBL_STPLC RIBULOSE BISPHOSPHATE | 1,058-01 |
| 35 | 60 | 42.0 | 459 | 1 | RBL_NYPER RIBULOSE BISPHOSPHATE | 1,058-01 |
| 36 | 60 | 42.0 | 465 | 1 | RBL_PASGU RIBULOSE BISPHOSPHATE | 1,058-01 |
| 37 | 60 | 42.0 | 465 | 1 | RBL_CAS21 RIBULOSE BISPHOSPHATE | 1,058-01 |
| 38 | 60 | 42.0 | 465 | 1 | BL_NECCA RIBULOSE BISPHOSPHATE | 1,058-01 |
| 39 | 60 | 42.0 | 467 | 1 | RBL_CALUS RIBULOSE BISPHOSPHATE | 1,058-01 |
| 40 | 60 | 42.0 | 475 | 1 | RBL_MAGAC RIBULOSE BISPHOSPHATE | 1,058-01 |
| 41 | 60 | 42.0 | 475 | 1 | RBL_PINRA RIBULOSE BISPHOSPHATE | 1,058-01 |
| 42 | 60 | 42.0 | 475 | 1 | RBL_PINTA RIBULOSE BISPHOSPHATE | 1,058-01 |
| 43 | 60 | 42.0 | 475 | 1 | RBL_KETDA RIBULOSE BISPHOSPHATE | 1,058-01 |
| 44 | 60 | 42.0 | 477 | 1 | RBL_CICIN RIBULOSE BISPHOSPHATE | 1,058-01 |
| 45 | 60 | 42.0 | 479 | 1 | RBL_BRAOL RIBULOSE BISPHOSPHATE | 1,058-01 |

ALIGNMENTS

RESULT 1
 ID SBP_CRYJA STANDARD: PRT: 374 AA.
 AC P18632:
 DI 01-NOV-1990 (Rel. 16, Create
 DI 01-NOV-1995 (Rel. 32, Last se-ence update)
 DE SUGI BASIC PROTEIN PRECURSOR (SBP) (MAJOR ALLERGEN CRY 3.1) (CRY 3.1)
 OS Cryptomeria japonica (Japanese cedar)
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC eukaryophytes; Spermatophytes; Coniferopsida; Coniferales;
 OC Taxodiaceae; Cryptomeria;
 RN [1]
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 RC TISSUE:POLLEN;
 RX MEDLINE: 94183234
 RA Sone T., Komiyama N., Shimizu K., Kusakabe T., Morikubo K.,
 RA Kino K.,
 RT Cloning and sequencing of cDNA coding for Cry 3.1, a major allergen
 RT of Japanese cedar pollen.
 RL Biochem. Biophys. Res. Commun. 199:519-525(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE:POLLEN;
 RA Namba M., Kato M., Iorihata K., Fushida S., Kurimoto M.,
 RN Submitted (JUL-1994) to the EMBL/Genbank/DBJ databases
 RP SEQUENCE OF 22-411
 RC TISSUE:POLLEN;
 RX MEDLINE: 89031257
 RA Tanai M., Ando S., Usui M., Kurimoto M., Sakaguchi M., Inouye S.,
 RA Matsuhashi T.,
 RT N-terminal amino acid sequence of a major allergen of Japanese cedar
 RT pollen (Cry 3.1).
 RL FEBS Lett. 239:329-332(1988).
 RN [4]
 RP CARBOHYDRATES.
 RX MEDLINE: 95003748
 RA Hijioka A., Matsumoto I., Kojima K., Ogawa H.,
 RA Antigenicity of the oligosaccharide moiety of the Japanese cedar
 RT (Cryptomeria japonica) pollen allergen, Cry 3.1.
 RL Int. Arch. Allergy Immunol. 105:198-202(1994).
 CC -!- PTN: CONTAINS FUCOSE/XYLOSE-CONTAINING N-LINKED OLIGOSACCHARIDES.
 CC -!- DISEASE: THIS PROTEIN IS THE MAJOR ALLERGEN OF JAPANESE CEDAR
 CC POLLEN. THE MOST COMMON POLLEN ALLERGEN IN JAPAN
 CC -!- MISCELLANEOUS: THE SEQUENCE OF CRY 3.1 FORM A IS SHOWN HERE. FORM
 CC B DIFFERS IN SIX POSITIONS.
 CC -!- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.


```

DR SEQ: L0004031: AOS1.
DR PFAM: PF003895: ThiF_family; 1.
KW DNA damage.
SC SEQUENCE 347 AA: 39273 MW: 1AD37B817BA27F3E CRC64:

Query Match 51.7% Score 74: DR 1: Length 347:
Best Local Similarity 52.9% Pred. No. 3.49e-02:
Matches 5: Conservative 5: Mismatches 3: Indels 0: Gaps 0:

Dr 140 NTLKRLNIPLYVAGSN 156
: 1 1111111111
QY 4 KNLNKLKMPLYIAGNK 20

RESULT 4
ID MP14_AYBAP STANDARD: PRT: 392 AA.
AC P28749:
D: 01-DEC-1992 (Rel. 24, Created)
D: 01-DEC-1992 (Rel. 24, Last sequence update)
DE POLLEN ALLERGEN AMB A 1.4 PRECURSOR (ANTIGEN 5) (ANTIGEN AMB A 1).
OS Ambrosia artemisiifolia (Short ragweed).
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
OC euphyllophytes: Spermatophyta: Magnoliophyta: eudicotyledons:
OC core eudicots: Asterales: euasterids II: Asterales: Asteraceae:
OC Asteroideae: Heliantheae: Ambrosia.
CC [1]
RN SEQUENCE FROM N.A., AND VARIANTS.
RP TISSUE: POLLEN.
RC MEDLINE: 52234570.
RA Griffith J.J., Pollock J., Klapper D.G., Rogers B.L., Nault A.K.:
"sequence polymorphism of Amb A 1 and Amb A 11, the major allergens
RI in Ambrosia artemisiifolia (Short ragweed).";
RE Int. Arch. Allergy Appl. Immunol. 96:245-304(1997).
CC -1- SUBUNIT: MONOMER.
CC -1- TISSUE SPECIFICITY: POLLEN AND FLOWERS.
CC -1- DISEASE: ONE OF THE MAJOR ALLERGENS OF THE RAGWEED POLLEN.
CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
CC AMB A 1/AMB A 11/CRY J 1 SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).

EMBL: M8C562; AAA32670.1
PIR: D53240; D53240.
PFAM: PF003844; pec_lyase_1.
PRINTS: PR00607; AMBALLERGEN.
Antigen: Allergen; Signal: Multigene family; Polymorphism.
FT SIGNAL 1 25
FT CHAIN 26 392 POLLEN ALLERGEN AMB A 1.4.
FT MOD_RES 226 226 BLOCKED.
FT VARIANT 182 188 SHGPPV -> CNDGPPA.
FI VARIANT 182 188
SQ SEQUENCE 392 AA: 42842 MW: 7B6219C12F365DA9 CRC64:

Query Match 46.9% Score 67: DB 1: Length 392:
Best Local Similarity 35.0% Pred. No. 6.68e-01:
Matches 7: Conservative 7: Mismatches 6: Indels 0: Gaps 0:

Dr 119 IFADWVIRDLRELAINDK 138
: 1 1111111111
QY 1 IFKNLNKLKMPLYIAGNK 20

RESULT 5
ID PELLILLO STANDARD: PRT: 434 AA.
AC P40973:
D: 01-FEB-1995 (Rel. 31, Created)
D: 01-FEB-1995 (Rel. 31, Last sequence update)
DE POLLEN ALLERGEN AMB A 1.4 PRECURSOR (ANTIGEN 5) (ANTIGEN AMB A 1).
OS Ambrosia artemisiifolia (Short ragweed).
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
OC euphyllophytes: Spermatophyta: Magnoliophyta: eudicotyledons:
OC core eudicots: Asterales: euasterids II: Asterales: Asteraceae:
OC Asteroideae: Heliantheae: Ambrosia.
CC [1]
RN SEQUENCE FROM N.A., AND VARIANTS.
RP TISSUE: POLLEN.
RC MEDLINE: 52234570.
RA Griffith J.J., Pollock J., Klapper D.G., Rogers B.L., Nault A.K.:
"sequence polymorphism of Amb A 1 and Amb A 11, the major allergens
RI in Ambrosia artemisiifolia (Short ragweed).";
RE Int. Arch. Allergy Appl. Immunol. 96:245-304(1997).
CC -1- SUBUNIT: MONOMER.
CC -1- TISSUE SPECIFICITY: POLLEN AND FLOWERS.
CC -1- DISEASE: ONE OF THE MAJOR ALLERGENS OF THE RAGWEED POLLEN.
CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
CC AMB A 1/AMB A 11/CRY J 1 SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).

EMBL: M8C562; AAA32670.1
PIR: D53240; D53240.
PFAM: PF003844; pec_lyase_1.
PRINTS: PR00607; AMBALLERGEN.
Antigen: Allergen; Signal: Multigene family; Polymorphism.
FT SIGNAL 1 25
FT CHAIN 26 392 POLLEN ALLERGEN AMB A 1.4.
FT MOD_RES 226 226 BLOCKED.
FT VARIANT 182 188 SHGPPV -> CNDGPPA.
FI VARIANT 182 188
SQ SEQUENCE 392 AA: 42842 MW: 7B6219C12F365DA9 CRC64:

Query Match 45.2% Score 66: DB 1: Length 434:
Best Local Similarity 43.0% Pred. No. 1.00e-00:
Matches 8: Conservative 6: Mismatches 6: Indels 0: Gaps 0:

Dr 156 IFGKSNVIRLKLQELINNK 175
: 1 1111111111
QY 1 IFKNLNKLKMPLYIAGNK 20

RESULT 6
ID V144_METIA STANDARD: PRT: 449 AA.
AC G57428:
D: 01-NOV-1997 (Rel. 35, Created)
D: 01-NOV-1997 (Rel. 35, Last sequence update)
D: 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPO-RETICAL PROTEIN M30164.
GN M30164.
OS Methanococcus jannaschii.
OC Archaea: Euryarchaeota: Methanococcales: Methanococcaceae:
CC Methanococcus.
CC [1]
RN SEQUENCE FROM N.A.
RP STRAIN=JAL-1 / DSM 2661 / ATCC 43067:
RX MEDLINE: 96337999.
RA Bait C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kleravage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.E., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cottor M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.:
"Complete genome sequence of the methanogenic archaeon, Methanococcus
RI jannaschii.";
RL Science 273:1056-1073(1996).
CC

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outposts on
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch.)
 CC -----
 DR EMBL: U67473; AAB98152.1;
 DR IGR: M0264;
 DR SWISS-PROT: P000157; RUBISCO_LARGE.1;
 KW Hypothetical protein.
 SQ SEQUENCE 395 AA: 46273 MW: 5E4AD0ABCC9A412 CRC64:

Query Match 44.8% Score 647 DB 1: Length 395:
 Best Local Similarity 64.3% Pred No. 2.23e-003
 Matches 9: Conservative 2: Mismatches 3: Indels 0: Gaps 0:

Db 129 IVEKYLNIKUD:PL 142
 QY 1 IFSKNEN:KLNMPYIAG 16

RESULT 2
 ID RBL_RUBS1 STANDARD: FBI: 453 AA.

AC Q30501: 1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE RUBULOSE BISPHOSPHATE CARBOXYLASE LARGE CHAIN PRECURSOR (EC 4.1.1.35)
 CE (RUBISCO LARGE SUBUNIT) (FRAGMENT).
 GN RUB1.
 OS Rubia tinctorum (Madder).
 OC Chloroplast.
 OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
 OC euphyllophytes: Spermatophyta: Magnoliophyta: eudicotyledons:
 OC core eudicots: Asteridiales: euasterids I: Gentianales: Rubiaceae:
 OC Rubia.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE: 9613925.
 RA Manc J.F., Nacali A.
 RI "Comparison of the evolution of ribulose 1,5-bisphosphate carboxylase
 RI (rbcL) and atp-rbcL noncoding spacer sequences in a recent plant
 RI group, the tribe Rubieae (Rubiaceae)." J. Mol. Evol. 41:920-927(1995)
 RL J. Mol. Evol. 41:920-927(1995)
 CC -1- FUNCTION: RUBISCO CATALYZES TWO REACTIONS: THE CARBOXYLATION OF
 CC RUBULOSE 1,5-BISPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC
 CC CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF
 CC THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH
 CC REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME
 CC ACTIVE SITE.
 CC -1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + CO(2) ->
 CC 2 3-PHOSPHO-D-GLYCERATE.
 CC -1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + O(2) ->
 CC 3-PHOSPHO-D-GLYCERATE + 2-PHOSPHOGLYCOLATE.
 CC -1- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS.
 CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.
 CC -1- SIMILARITY: BELONGS TO THE RUBISCO LARGE CHAIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outposts on
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch.)
 CC -----

DR EMBL: X81104; CAA57010.1;
 DR HSSP: P00876; RBL.
 DR MENDEL: 4673; RUBT1; rbcL.1.
 DR PROSITE: P500157; RUBISCO_LARGE.1.
 DR PFAM: PF00016; RuBisCO_large.1.
 KW Photosynthesis: Carbon dioxide fixation: Photorespiration:

KW Lyase: Oxidoreductase: Monooxygenase: Chloroplast: Acetylation.
 FT PROPEP 1 2
 FT CHAIN 3 >453
 FT MOD_RES 3 3
 FT ACT_SITE 201 201
 FT NON_TER 453 453
 SQ SEQUENCE 453 AA: 50263 MW: 88853FEA72F849A0 CRC64:

Query Match 44.8% Score 647 DB 1: Length 453:
 Best Local Similarity 33.3% Pred No. 3.4e-003
 Matches 6: Conservative 7: Mismatches 5: Indels 0: Gaps 0:

Db 255 VFARLAVPIVMEHYTG 272
 QY 1 IFSKNEN:KLNMPYIAG 16

RESULT 8
 ID RBL_LUPM1 STANDARD: FBI: 455 AA.

AC P92406:
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE RUBULOSE BISPHOSPHATE CARBOXYLASE LARGE CHAIN (EC 4.1.1.39) (RUBISCO
 DE LARGE SUBUNIT) (FRAGMENT).
 GN RUB1.
 OS Lupinus microcarpus.
 OC Chloroplast.
 OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
 OC euphyllophytes: Spermatophyta: Magnoliophyta: eudicotyledons:
 OC core eudicots: Rosidae: eurosids I: Fabales: Fabaceae: Papilionoideae:
 OC Lupinus.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA TISSUE:LEAF.
 RA Kaess E., Wink M.
 RI "Molecular phylogeny of the Papilionaceae (family Leguminosae): first
 RI sequences versus chemical taxonomy." Bot. Acta 108:149-162(1995).
 CC -1- FUNCTION: RUBISCO CATALYZES TWO REACTIONS: THE CARBOXYLATION OF
 CC D-RIBULOSE 1,5-BISPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC
 CC CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF
 CC THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH
 CC REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME
 CC ACTIVE SITE.
 CC -1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + CO(2) ->
 CC 2 3-PHOSPHO-D-GLYCERATE.
 CC -1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + O(2) ->
 CC 3-PHOSPHO-D-GLYCERATE + 2-PHOSPHOGLYCOLATE.
 CC -1- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS.
 CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.
 CC -1- SIMILARITY: BELONGS TO THE RUBISCO LARGE CHAIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outposts on
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch.)
 CC -----

DR EMBL: Z70063; CAA93922.1;
 DR PROSITE: P500157; RUBISCO_LARGE.1.
 DR PFAM: PF00016; RuBisCO_large.1.
 KW Photosynthesis: Carbon dioxide fixation: Photorespiration:
 KW Lyase: Oxidoreductase: Monooxygenase: Chloroplast.
 FT ACT_SITE 192 192
 FT NON_TER 455 455
 SQ SEQUENCE 455 AA: 50265 MW: 4610CF5DF6145933 CRC64:

Query Match 44.8% Score 63: DB 1: Length 455:

Best Local Similarity 33.3%; Pred. No. 3.31e+00; Mismatches 7; Indels 0; Gaps 0;

DB 246 VFARELGVPVIMHDIYTG 263
 QY 1 IFSKNLNKLNPLYIAG 18

RESULT 9

ID RBL_CERGL STANDARD; PRI: 475 AA.
 AC P25830;
 DT 01-MAY-1992 (Rel. 22, Created);
 DT 01-MAY-1992 (Rel. 22, Last sequence update);
 DT 15-DEC-1998 (Rel. 37, Last annotation update);
 DE RIBULOSE BISPHOSPHATE CARBOXYLASE LARGE CHAIN PRECURSOR (EC 4.1.1.39)
 DE (RUBISCO LARGE SUBUNIT);
 GN RBCL;
 OS Cerastium glomeratum;
 OC Chloroplast;
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 OC core eudicots; Caryophyllales; Caryophyllaceae;
 OC Cerastium;
 RN [1]

SEQUENCE FROM N.A.
 RA Manhart J.R., Hugh J.H., Wilson D.;
 RL Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: RUBISCO CATALYZES TWO REACTIONS: THE CARBOXYLATION OF
 CC D-RIBULOSE 1,5-BISPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC
 CC CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF
 CC THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH
 CC REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME
 CC ACTIVE SITE.
 CC -!- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + CO(2) =
 CC 2 3-PHOSPHO-D-GLYCERATE;
 CC -!- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + O(2) =
 CC 3-PHOSPHO-D-GLYCERATE + 2-PHOSPHOGLYCOLATE;
 CC -!- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS.
 CC -!- SUBCELLULAR LOCATION: CHLOROPLAST;
 CC -!- SIMILARITY: BELONGS TO THE RUBISCO LARGE CHAIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M83542; AAA84173; 1;
 CC RSP: P00875; 1AUS;
 CC MENDEL: 2754; CERGL:rbcl.1;
 CC PROSITE: PS00157; RUBISCO_LARGE; 1;
 CC PFAM: PF00016; RUBISCO_LARGE; 1;
 CC Photosynthesis; Carbon dioxide fixation; Photorespiration;
 CC Lyase; Oxidoreductase; Monooxygenase; Chloroplast; Acetylation.
 CC PROPEP 1 2
 CC CHAIN 3 475
 CC RIBULOSE BISPHOSPHATE CARBOXYLASE LARGE
 CC CHAIN
 CC MOD_RES 3 3
 CC ACT_SITE 201 201
 CC BINDING OF CO(2) ACTIVATES THE ENZYME.
 CC SEQUENCE 475 AA: 52656 MW: 90986466 E0211B CRC64;

Query Match 44.18; Score 63; DB 1; Length 475;
 Best Local Similarity 33.3%; Pred. No. 3.31e+00;
 Mismatches 6; Conservative 7; Indels 0; Gaps 0;

DB 255 VFARELGVPVIMHDIYTG 272
 QY 1 IFSKNLNKLNPLYIAG 18

RESULT 10

ID RBL_CERGL STANDARD; PRI: 476 AA.
 AC P25827;
 DT 01-MAY-1992 (Rel. 22, Created);
 DT 01-MAY-1992 (Rel. 22, Last sequence update);
 DT 15-DEC-1998 (Rel. 37, Last annotation update);
 DE RIBULOSE BISPHOSPHATE CARBOXYLASE LARGE CHAIN PRECURSOR (EC 4.1.1.39)
 DE (RUBISCO LARGE SUBUNIT);
 GN RBCL;
 OS Arenaria drummondii;

Best Local Similarity 33.3%; Pred. No. 3.31e+00; Mismatches 7; Indels 0; Gaps 0;

DB 246 VFARELGVPVIMHDIYTG 263
 QY 1 IFSKNLNKLNPLYIAG 18

RESULT 9

ID RBL_STEME STANDARD; PRI: 475 AA.
 AC P46920;
 DT 01-NOV-1995 (Rel. 32, Created);
 DT 01-NOV-1995 (Rel. 32, Last sequence update);
 DT 15-DEC-1998 (Rel. 37, Last annotation update);
 DE RIBULOSE BISPHOSPHATE CARBOXYLASE LARGE CHAIN PRECURSOR (EC 4.1.1.39)
 DE (RUBISCO LARGE SUBUNIT);
 GN RBCL;
 OS Stelleria media;
 OC Chloroplast;
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 OC core eudicots; Caryophyllales; Caryophyllaceae;
 OC Stelleria;
 RN [1]

SEQUENCE FROM N.A.
 RA Rettig J.H., Wilson H.D., Manhart J.R.;
 RL Submitted (XXX-1994) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: RUBISCO CATALYZES TWO REACTIONS: THE CARBOXYLATION OF
 CC D-RIBULOSE 1,5-BISPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC
 CC CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF
 CC THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH
 CC REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME
 CC ACTIVE SITE.
 CC -!- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + CO(2) =
 CC 2 3-PHOSPHO-D-GLYCERATE;
 CC -!- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + O(2) =
 CC 3-PHOSPHO-D-GLYCERATE + 2-PHOSPHOGLYCOLATE;
 CC -!- SUBUNIT: 8 LARGE CHAINS + 3 SMALL CHAINS.
 CC -!- SUBCELLULAR LOCATION: CHLOROPLAST;
 CC -!- SIMILARITY: BELONGS TO THE RUBISCO LARGE CHAIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M62570; AAA84620.1;
 CC RSP: P00875; 1AUS;
 CC MENDEL: 2993; STEME:rbcl.1;
 CC PROSITE: PS00157; RUBISCO_LARGE; 1;
 CC PFAM: PF00016; RUBISCO_LARGE; 1;
 CC Photosynthesis; Carbon dioxide fixation; Photorespiration;
 CC Lyase; Oxidoreductase; Monooxygenase; Chloroplast; Acetylation.
 CC PROPEP 1 2
 CC CHAIN 3 475
 CC RIBULOSE BISPHOSPHATE CARBOXYLASE LARGE
 CC CHAIN
 CC MOD_RES 3 3
 CC ACT_SITE 201 201
 CC BINDING OF CO(2) ACTIVATES THE ENZYME.
 CC SEQUENCE 475 AA: 52675 MW: 90455367 E0785F1 CRC64;

Query Match 44.18; Score 63; DB 1; Length 475;
 Best Local Similarity 33.3%; Pred. No. 3.31e+00;
 Mismatches 6; Conservative 7; Indels 0; Gaps 0;

DB 255 VFARELGVPVIMHDIYTG 272
 QY 1 IFSKNLNKLNPLYIAG 18

RESULT 11

ID RBL_AREDR STANDARD; PRI: 476 AA.
 AC P25827;
 DT 01-MAY-1992 (Rel. 22, Created);
 DT 01-MAY-1992 (Rel. 22, Last sequence update);
 DT 15-DEC-1998 (Rel. 37, Last annotation update);
 DE RIBULOSE BISPHOSPHATE CARBOXYLASE LARGE CHAIN PRECURSOR (EC 4.1.1.39)
 DE (RUBISCO LARGE SUBUNIT);
 GN RBCL;
 OS Arenaria drummondii;

CC Chloroplast.
 CC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;
 CC euphyllophytes: Spermatophyta: Magnoliophyta: eudicotyledons:
 CC core eudicots: Caryophyllales: Caryophyllaceae: Caryophyllaceae;
 CC Arenaria.
 CC [1]
 RN SEQUENCE FROM N.A.
 RA Manhart J.R., Huth J.H., Wilson D.J.
 RC Submitted (XXX-1992) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: RUBISCO CATALYZES TWO REACTIONS: THE CARBOXYLATION OF
 CC D-RIBULOSE 1,5-BISPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC
 CC CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF
 CC THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH
 CC REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME
 CC ACTIVE SITE.
 CC -!- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + CO(2) =
 CC 2 3-PHOSPHO-D-GLYCERATE.
 CC -!- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + O(2) =
 CC 3-PHOSPHO-D-GLYCERATE + 2-PHOSPHOGLYCOLATE.
 CC -!- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS.
 CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.
 CC -!- SIMILARITY: BELONGS TO THE RUBISCO LARGE CHAIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: M35421; AAA84023.1;
 DR HSP: P00875; LAUS.
 DR MENDEL: 2732; ARED; rbcL1.
 DR PROSITE: PS00157; RUBISCO_LARGE; 1.
 DR PFAM: PF00016; RuBisCO_large; 1.
 KW Photosynthesis; Carbon dioxide fixation; Photorespiration;
 KW Lyase; Oxidoreductase; Monooxygenase; Chloroplast; Acetylation.
 FT PROPEP 1 2
 FT CHAIN 3 476
 FT MOD_RES 3 3
 FT ACT_SITE 201 201
 FT SEQUENCE 476 AA; 52662 MW; B4B913047B81589 CRC64;
 Query Match 44.1%; Score 63; DB 1; Length 476;
 Best Local Similarity 33.3%; Pred. No. 3,31e-00;
 Matches 6; Conservative 7; Mismatches 5; Indels 0; Gaps 0;
 DB 255 VFARELGVPIVNHDIYIG 272
 QY 1 IFSKNNLKNMPLYIAG 18
 RESULT 12
 ID RBLAVESA STANDARD: PRT: 477 AA.
 AC P48684;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 01-DEC-1998 (Rel. 37, Last annotation update)
 DE RIBULOSE BISPHOSPHATE CARBOXYLASE LARGE CHAIN PRECURSOR (EC 4.1.1.39)
 DE (RUBISCO LARGE SUBUNIT).
 GN RBCL.
 OS Avena sativa (Oat).
 OG Chloroplast.
 CC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;
 CC euphyllophytes: Spermatophyta: Magnoliophyta: Liliopsida: Poales;
 CC Poaceae: Avena.
 CC [1]
 RN SEQUENCE FROM N.A.
 RA Garcia P.G.
 RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: RUBISCO CATALYZES TWO REACTIONS: THE CARBOXYLATION OF

CC D-RIBULOSE 1,5-BISPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC
 CC CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF
 CC THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH
 CC REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME
 CC ACTIVE SITE.
 CC -!- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + CO(2) =
 CC 2 3-PHOSPHO-D-GLYCERATE.
 CC -!- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + O(2) =
 CC 3-PHOSPHO-D-GLYCERATE + 2-PHOSPHOGLYCOLATE.
 CC -!- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS.
 CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.
 CC -!- SIMILARITY: BELONGS TO THE RUBISCO LARGE CHAIN FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: L15300; AAA84028.1;
 DR HSP: P00876; 3RUB.
 DR MENDEL: 2735; AVEAR; rbcL1.
 DR PROSITE: PS00157; RUBISCO_LARGE; 1.
 DR PFAM: PF00016; RuBisCO_large; 1.
 KW Photosynthesis; Carbon dioxide fixation; Photorespiration;
 KW Lyase; Oxidoreductase; Monooxygenase; Chloroplast; Acetylation.
 FT PROPEP 1 2
 FT CHAIN 3 477
 FT MOD_RES 3 3
 FT ACT_SITE 201 201
 FT SEQUENCE 477 AA; 52935 MW; F1D48ED95AED3DD CRC64;
 Query Match 44.1%; Score 63; DB 1; Length 477;
 Best Local Similarity 33.3%; Pred. No. 3,31e-00;
 Matches 6; Conservative 7; Mismatches 5; Indels 0; Gaps 0;
 DB 255 VFARELGVPIVNHDIYIG 272
 QY 1 IFSKNNLKNMPLYIAG 18
 RESULT 13
 ID Y454_MYCGE STANDARD: PRT: 155 AA.
 AC P47692;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE MYCETHELICAL PROTEIN M454
 GN M454
 OS Mycoplasma genitalium.
 CC Bacteria: Firmicutes: Bacillus/Clostridium group: Mollicutes
 CC Mycoplasmataceae: Mycoplasma.
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN:ATCC 33530 / G-37;
 RX MEDLINE: 96026345.
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
 RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
 RA Fritchman J.L., Weidman J.F., Sall K.V., Sandusky M., Fuhrman J.L.,
 RA Nguyen D.T., Utterback T.R., Saiki H.K., Phillips C.A., Merrick J.M.,
 RA Tomb J.-F., Dougherty B.A., Boff K.F., Hu P.-C., Lucier T.S.,
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
 RT "The minimal gene complement of Mycoplasma genitalium."
 RC Science 270:397-403(1995).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial


```

RESULT 15
ID TP6A_SULSH STANDARD PRI 389 AA
AC O05208
DT 15-JUL-1998 (Rel. 36, Created)
DI 15-JUL-1998 (Rel. 36, Last sequence update)
OI 15-JUL-1998 (Rel. 36, Last annotation update)
DE TYPE 1: DNA TOPOISOMERASE VI SUBUNIT A (EC 5.99.1.3)
CN TOP6A
OS Sulfolobus shibatae
OC Archaea: Crenarchaeota: Sulfolobales: Sulfolobus
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 212-221 AND 344-352.
RC STRAIN-B12
RX MEDLINE: 97218655
RA Bergerat A., de Massy B., Gadelle D., Voronias P.-C., Nicolas A.,
  Forterre P.
RI *An atypical topoisomerase II from Archaea with implications for
  meiotic recombination.
RL Nature 386:414-417(1997).
CC -- FUNCTION: RELAXES BOTH POSITIVE AND NEGATIVE SUPERTEURNS AND
  EXHIBITS A STRONG DECATENASE ACTIVITY.
CC -- CATALYTIC ACTIVITY: ATP-DEPENDENT; BREAKAGE, PASSAGE AND REJOINING
  OF DOUBLE-STRANDED DNA.
CC -- SUBUNIT: HETEROTETRAMER OF TWO SUBUNITS A AND TWO SUBUNITS B.
CC -- SIMILARITY: BELONGS TO THE TOP6A FAMILY.
CC -----
CC This SWISS-prot entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
  the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y10582; CAA71635.1;
KW Isomerase; Topoisomerase; DNA-binding;
FT ACT_SITE 106-106 DNA_CLEAVAGE (BY SIMILARITY);
SQ SEQUENCE 389 AA; 45055 MW; 6F617CB8231431B7 CRC32;

```

```

Query Match 43.4%; Score 62; DB 1; Length 389;
Best local similarity 46.7%; Pred. No. 4; E-00;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

```

```

DB 244 FVRRLNEELKLPVYI 258
QY 1 11 11:11
   2 FSKNLNKLNPVLYI 16

```

```

Search completed: Mon Jun 19 16:13:01 2000
Job time : 5 secs.

```


DE POLLEN MAJOR ALLERGEN 1-2.
 OS Juniperus ashei (Ozark white cedar).
 OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
 euphyllophytes: Spermatophyta: Coniferopsida: Coniferales:
 OC Taxodiaceae: Juniperus.
 RN (1)
 RP SEQUENCE FROM N.A.
 RA MIDORI HORIUTI I.M., GOLDBLUM R.M., KUROSKY A., WOOD T.G.,
 RA BROOKS E.G.,
 WJ "Molecular cloning of mountain cedar (Juniperus ashei) pollen major
 RI allergen, Jun a 1".
 RI Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
 CR EMBL: AF105663; AAC03609.1;
 CR EMBL: AF105662; AAC03608.1;
 CR MENDEL: 36544; Juncas:1388;36544
 CR MENDEL: 36545; Juncas:1388;36545
 CR MENDEL: 36545; Juncas:1388;36545
 SQ SEQUENCE 367 AA; 5924 MW; 472UB692 CF 32;

Query Match 86.0%; Score 123; DB 10; Length 367;
 Best Local Similarity 75.0%; Pred. No. 1,470-10;
 Matches 15; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Db 92 IFSQNM:IKKMPLYVAGK 112

Qy 1 IFSKNLNKLNMPLYAGNK 20

RESULT 3
 ID Q98879 PRELIMINARY: PRT: 444 AA.

AC Q98879
 DT 01-MAY-1999 (TrEMBLrel: 10, Created)

DT 01-MAY-1999 (TrEMBLrel: 10, Last sequence update)

DT 01-NOV-1999 (TrEMBLrel: 12, Last annotation update)

DE RIBULOSE BIPHOSPHATE CARBOXYLASE LARGE CHAIN (EC 4.1.1.39)

DE (FRAGMENT).

GN RBCL

OS Couratari tatarica

OC Chloroplast

OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:

euphyllophytes: Spermatophyta: Magnoliophyta: eudicotyledons:

core eudicots: Asteridae: Ericales: Lecythidaceae: Couratari.

RN (1)

RP SEQUENCE FROM N.A.

RA MORTON C.M., MORI S.A., PRANCE G.T., KAROL K.G., CHASE M.W.;

*Phylogenetic relationships of Lecythidaceae: A cladistic analysis

using rbcL sequence and morphological data";

RA Am. J. Bot. 84:530-540(1997).

CC -1- FUNCTION: RUBISCO CATALYSES TWO REACTIONS: THE CARBOXYLATION OF D-

RIBULOSE 1,5-BIPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHESIS;

CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF

THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATORY PATHWAY. BOTH

REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME

ACTIVE SITE.

CC -1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BIPHOSPHATE + CO(2) = 2 3-

PHOSPHO-D-GLYCERATE.

CC -1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BIPHOSPHATE + O(2) = 3-

PHOSPHO-D-GLYCERATE + 2-PHOSPHOGLYCOLATE.

CC -1- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS.

CR EMBL: 230177; CAB0224.1;

CR MENDEL: 36884; Couratari:36884.

CR PROSITE: PS00157; RUBISCO_LARGE.1;

KW Chloroplast; Photosynthesis; Carbon dioxide fixation;

Photorepiration; Lyase; Oxidoreductase; Monooxygenase.

FT NON-TER 1

FT ACT_SITE 444

FT ACT_SITE 178

SQ SEQUENCE 444 AA; 43296 MW; C98E6781 CRC32;

Query Match 47.68; Score 68; DB 8; Length 444;
 Best Local Similarity 27.8%; Pred. No. 3,048-00;
 Matches 5; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Db 231 VFARELGPXMHXYLTG 243

Qy 1 IFSKNLNKLNMPLYAG 18

RESULT 4
 ID Q64510 PRELIMINARY: PRT: 455 AA.

AC Q64510

DT 01-AUG-1998 (TrEMBLrel: 07, Created)

DT 01-AUG-1999 (TrEMBLrel: 07, Last sequence update)

DT 01-NOV-1999 (TrEMBLrel: 12, Last annotation update)

DE PUTATIVE PEPTIDE HYDROLASE

GN T20F6.14.

OS Arabidopsis thaliana (Mouse-ear cress);

OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:

euphyllophytes: Spermatophyta: Magnoliophyta: eudicotyledons:

core eudicots: Rosidae: eurosids 1: Brassicales: Brassicaceae:

OC Arabidopsis.

RN (1)

RP SEQUENCE FROM N.A.

RA STRAIN-OV; COLUMBIA;

RA ROUNSELEY S.D., LIN X., KETCHUM K.A., CRUSHY M.L., BRAND-N W.D.,

RA SYKES S.M., KAUL S., WASON T.M., KERLAVALA A.R., ADAMS M.C.,

RA SOMERVILLE C.R., VENTER J.C.,

RA Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases

CR EMBL: AC002521; AAC05950.1;

CR MENDEL: 28505; Arabid:1088;28505.

CR PFAM: PF00544; pec_lyase.1.

CR PRINTS: PRC0807; AMBALLERGEN.

KW Lyase

SQ SEQUENCE 455 AA; 51257 MW; E5ACH62D CRC32;

Query Match 47.6%; Score 68; DB 10; Length 455;
 Best Local Similarity 40.0%; Pred. No. 3,048-00;
 Matches 8; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Db 171 IFARSMIKLOCELIITNCK 19

Qy 1 IFSKNLNKLNMPLYAGNK 20

RESULT 5
 ID Q21303 PRELIMINARY: PRT: 1202 AA.

AC Q21303

DT 01-NOV-1996 (TrEMBLrel: 01, Created)

DT 01-NOV-1996 (TrEMBLrel: 01, Last sequence update)

DT 01-JAN-1999 (TrEMBLrel: 09, Last annotation update)

DE K07G5.3 PROTEIN.

GN K07G5.3

OS Caenorhabditis elegans.

OC Eukaryota: Metazoa: Nematoda: Secernentea: Rhabditia: Rhabditidae:

OC Rhabditina: Rhabditidae: Rhabditidae: Rhabditidae: Rhabditidae:

RN (1)

RP SEQUENCE FROM N.A.

RA MCMURRAY A.

RA Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.

RN (2)

RP SEQUENCE FROM N.A.

RA MEDLINE: 94150719.

RA WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M.,

RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COLLISON A.,

RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,

RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,

RA JONES M., KERSHAW J., KIRSTEN N., LAISTER N., LAKEILJE P.,

RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,

RA PARSONS J., PERCY C., RIFKIN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,

RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULTON J.,

RA THIERVY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,

RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WORMAN P.,

RA "2.2 Mb of contiguous nucleotide sequence from chromosome 11: of C.

RT elegans."

RL Nature 368:32-38(1994).

DR EMBL: 271264; GAA9830.1;

SQ SEQUENCE 1202 AA; 138469 MW; 6F9C9E53 CRC32;

Query Match 47.6% Score 68; DB 5; Length 1202;
 Best Local Similarity 51.5% Pred. No. 3,04e+00;
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 8:1 IRLNPLYISNKG 823
 :||:||||:
 Qy 8 IKLNPLYIAGNK 20

RESULT 6
 ID 078617 PRELIMINARY: PRI: 430 AA.
 AC 078617
 DT 01-NOV-1998 (TREMBlrel. 08, Created)
 DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 12, Last annotation update)
 DE RIBULOSE BIPHOSPHATE CARBOXYLASE LARGE CHAIN (EC 4.1.1.39)
 DE (FRAGMENT).
 GN RBCL.
 OS Chlorococc.
 OS Chloroplast.
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 eubryophytes; Spermatophyta; Magnoliophyta; Liliopsida; Asparagales;
 OC Orchidaceae; Meliracillium.
 RN 11

SEQUENCE FROM N.A.
 RA CAMERON K.M., CHASE M.W., WHITTEN W.M., KORES P.J., JARRELL D.C.,
 RA ALBERT V.A., YUKAWA T., HILLS H.G., GOLDMAN D.H.;
 RT "A phylogenetic analysis of the Orchidaceae: evidence from rbcL
 nucleotide sequences";
 RA Am. J. Bot. 0:0-0(1998);

!- FUNCTION: RUBISCO CATALYZES TWO REACTIONS: THE CARBOXYLATION OF D-
 RIBULOSE 1,5-BIPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC
 CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF
 THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH
 REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME
 ACTIVE SITE.
 !- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BIPHOSPHATE + CO(2) -> 2 3-
 PHOSPHO-D-GLYCERATE.
 !- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BIPHOSPHATE + O(2) -> 3-
 PHOSPHO-D-GLYCERATE + 2-PHOSPHOGLYCOLATE
 !- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS.
 EMBL: AF074192; AAC35168.1;
 HSP: P04718; GRUB.
 DR MENDEL: 30551; Meltr:rbcl:30551;
 DR PROSITE: PS00157; RUBISCO_LARGE; 1.
 DR PFAM: PF00016; RUBISCO_LARGE; 1.
 KW Chloroplast; Photosynthesis; Carbon dioxide fixation;
 KW Photorespiration; Lyase; Oxidoreductase; Monooxygenase.
 FT NON-TER 2 1
 FT NON-TER 430 430
 FT ACT_SITE 182 182
 FT ACT_SITE 182 182
 SQ SEQUENCE 430 AA: 47681 MW: 759E959E CRC32:

Query Match 45.9% Score 67; DB 8; Length 430;
 Best Local Similarity 33.3% Pred. No. 4.39e+00;
 Matches 6; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

Db 235 VFARELGVFIIMPDYLTG 252
 :||:||||:
 Qy 1 IFSKNLNKLNPLYIAG 18

RESULT 7
 ID 023665 PRELIMINARY: PRI: 459 AA.
 AC 023665
 DT 01-JAN-1998 (TREMBlrel. 05, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-NOV-1998 (TREMBlrel. 12, Last annotation update)
 DE PUTATIVE PECTATE LYASE.
 GN 59.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC eubryophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 OC core eudicots; Rosidae; eurosids 1; Brassicales; Brassicaceae;
 OC Arabidopsis.
 RN 11
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 97422403.
 RA KULIKASKAS R., MCCORMICK S.;
 RI "Identification of the tobacco and Arabidopsis homologues of the
 RI pollen-expressed LAT59 gene of tomato";
 RL Plant Mol. Biol. 34:809-814(1997).
 RL ENBL: U93619; AAB69759.1;
 DR MENDEL: 25559; Arabid088; 25559.
 DR PFAM: PF00544; pect_lyase; 1.
 DR PRINTS: PR00807; AMBALLERGEN.
 KW Lyase.
 SQ SEQUENCE 459 AA: 51420 MW: 41E9E908 CRC32:

Query Match 46.9% Score 67; DB 10; Length 459;
 Best Local Similarity 40.0% Pred. No. 4.39e+00;
 Matches 8; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

Db 175 IFARSMIKLQQLMITSXK 194
 :||:||||:
 Qy 1 IFSKNLNKLNPLYIAGNK 20

RESULT 8
 ID 018771 PRELIMINARY: PRI: 555 AA.
 AC 018771
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
 DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
 DE CS1E3.6 PROTEIN.
 GN CS1E3.6.
 OS Caenorhabditis elegans.
 OS Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;
 OC Rhabditina; Rhabditidae; Rhabditidae; Rhabditidae; Rhabditidae;
 RN 11
 RP SEQUENCE FROM N.A.
 RA WILKINSON J.;
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases
 [2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE: 94150718.
 RA WILSON R., AINSOUGH R., ANDERSON K., BAYNES C., BERKS M.,
 RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COLLINS N.A.,
 RA CRAXTON M., DEAR S., DU Z., DUNN R., FAVELLO A., FULTON L.,
 RA GARDNER A., GREEN P., HANKINS J., HILLIER L., JIER M., JOHNSTON L.,
 RA JONES M., KERSHAW J., KIRSTEN J., LAISER N., LAFFITTE P.,
 RA LIGHTNING J., LLOYD C., MENDRUP A., MORTIMER R., O'CALLAGHAN M.,
 RA PARSONS J., PERRY C., RIKEN L., POLERA A., SAUNDERS D., SHANKEN P.,
 RA SKALDON N., SMITH A., SUNDHAMMER E., STADEN R., SULSTON J.,
 RA THIERRY-MIEG J., THOMAS K., VAUGHAN M., VAUGHAN K., WATERSTON K.,
 RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT C., WOHLMAN P.,
 RT *2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RT elegans*.
 RL Nature 368:32-38(1994).
 DR ENBL: 278410; CAB01641.1;
 DR PFAM: PF00860; xan_ur_permease; 1.
 SQ SEQUENCE 555 AA: 60315 MW: 4D90DD42 CRC32:

Query Match 45.9% Score 67; DB 5; Length 555;
 Best Local Similarity 35.3% Pred. No. 4.39e+00;
 Matches 6; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

Db 464 FNQILNVLLQMPFVGA 480
 :||:||||:
 Qy 2 FSKNLNKLNPLYIAG 18

RESULT 9
 ID 031994 PRELIMINARY: PRI: 414 AA.
 AC 031994


```

DB      246 IFARELGVPIVHHDIITG 263
QY      : IFSKLNLIKMLPLYIAG 18

RESULT  12
ID       P93920 PRELIMINARY: PRI: 467 AA.
AC       P93920
CT       01-MAY-1997 (TEmBLrel_03, Created)
DT       01-MAY-1997 (TEmBLrel_03, Last sequence update)
DE       01-NOV-1999 (TEmBLrel_12, Last annotation update)
OE       RIBULOSE BISPHTOSPHATE CARBOXYLASE LARGE CHAIN (EC 4.1.1.39)
OF       (FRAGMENT).
OG       RBCL
OS       Odontostomum hartwegii.
OC       Chloroplast.
OO       Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OQ       euphyllphytes; Spermatophyta; Magnoliophyta; Liliopsida; Asparagales;
OC       Tecophilactaceae; Odontostomum.
OR       [1]
RP       SEQUENCE FROM N.A.
RC       TISSUE=LEAF;
RR       RUDALL P.J., FURNESS C.A., FAY M.F., CHASE M.W.:
RL       Submitted (JUL-1996) to the EMBL/GeneBank/DBS databases.
RO       -! FUNCTION: RUBISCO CATALYSES TWO REACTIONS: THE CARBOXYLATION OF D-
OC       RIBULOSE 1,5-BISPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC
CC       CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF
CCC      THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH
CCC      REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME
CC       ACTIVE SITE.
CC       -! CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + CO(2) = 2 3-
CC       PHOSPHO-D-GLYCERATE.
CC       -! CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + O(2) = 3-
CC       PHOSPHO-D-GLYCERATE + 2-PHOSPHOGLYCULATE.
CC       -! SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHA NS.
CC       -! SUBCELLULAR LOCATION: CHLOROPLAST.
CR       EMBL: Z77314; CABOL109.1;
DR       HSP: P00880; IRSC.
DE       MENDEL: 14408; OdoMa:rbcl:14408.
DD       PROSITE: PS00157; RUBISCO_LARGE: 1.
DK       PFAM: PF00016; RUBISCO_large: 1.
KW       Lyase; Chloroplast; Photosynthesis; Carbon dioxide fixation;
KK       Photorespiration; Oxidoreductase; Monooxygenase.
FT       NON_TER 1 192
FT       ACT_SITE 192 192 BINDING OF CO(2) ACTIVATES THE ENZYME.
FT       ACT_SITE 193 193 BINDING OF CO(2) ACTIVATES THE ENZYME.
FT       NON_TER 467 467
SQ       SEQUENCE 467 AA; 51835 MW; 1F12A356 CRC32;

Query Match 45.5% Score 65; Len 8; Length 467;
Best Local Similarity 27.8%; Prob. No. 9.38e+00;
Matches 5; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Dt      246 VFARELGVPVIVHHXYLTG 263
QY      : IFSKLNLIKMLPLYIAG 18

RESULT  13
ID       Q31807 PRELIMINARY: PRI: 475 AA.
AC       Q31807
DT       01-NOV-1996 (TEmBLrel_01, Created)
DT       01-NOV-1996 (TEmBLrel_01, Last sequence update)
DT       01-NOV-1999 (TEmBLrel_12, Last annotation update)
DE       RIBULOSE BISPHTOSPHATE CARBOXYLASE LARGE CHAIN (EC 4.1.1.39)
OF       (FRAGMENT).
OG       RBCL
OS       Angelica lucida.
OO       Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OQ       euphyllphytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC       core eudicots; Asteridae; euasterids II; Apiales; Apiaceae; Angelica.
OR       [1]

```

```

SEQUENCE FROM N.A.
PLUNKETT G.M., SOLTIS D.E., SOLTIS P.S.;
Am. J. Bot. 83:499-515 (1996).
-1- FUNCTION: RUBISCO CATALYZES TWO REACTIONS: THE CARBOXYLATION OF D-
RIBULOSE 1,5-BISPHOSPHATE THE PRIMARY EVENT IN PHOTOSYNTHETIC
CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF
THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH
REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME
ACTIVE SITE.
-1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + CO(2) -> D-3-
PHOSPHO-D-GLYCERATE.
-1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + O(2) ->
PHOSPHO-D-GLYCERATE + 2-PHOSPHOGLYCOLATE.
-1- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS.
-1- SUBCELLULAR LOCATION: CHLOROPLAST.
EMBL: U50221; AAB98935.1;
MENDEL: 21616; Anglu.rbcL216
PROSITE: PS00157; RUBISCO_LARGE
PFAM: PF00016; Rubisco_large
KW Chloroplast; Photosynthesis; Carbon dioxide fixation;
KW Photorespiration; Lyase; Oxidoreductase; Monooxygenase.
FT NON-TERMINAL
FT ACT-SITE 202 202 BINDING OF CO(2) ACTIVATES THE ENZYME
FT ACT-SITE 202 202 BINDING OF CO(2) ACTIVATES THE ENZYME
SQ SEQUENCE 475 AA: 52676 MW: 8593079 CRC32:
Query Match: 45.5%; Score 65; DB 8; Length 475;
Best Local Similarity 36.9%; Pred No. 9.08e+00;
Matches 7; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

Db 255 IFARELGVPVIMHDYITG 272
QY 1 IFSKNLKLNMPLYAG 18
|||||:|:|:|:|:|:|

RESULT 14
ID P92302 PRELIMINARY: PRT: 475 AA.
AC
D1 MAY-1997 (TREMBLER) C3. Created;
D1 MAY-1997 (TREMBLER) C3. Last sequence update;
D1 NOV-1999 (TREMBLER) 12. Last annotation update;
DE RIBULOSE BISPHOSPHATE CARBOXYLASE LARGE CHAIN (EC 4.1.1.39).
GN RBC1.
OS Hillebrandia sandwicensis.
OC Chloroplast.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Euphyllophytes; Spermatophytes; Magnoliophyta; eudicotyledons;
OC Core eudicots; Rosidae; eucots 1; Cucurbitales; Begoniaceae.
OC Hillebrandia.
RN
RA SWENSEN S.M.;
RA SEQUENCE FROM N.A.
RL Am. J. Bot. 83:503-512 (1996).
-1- FUNCTION: RUBISCO CATALYZES TWO REACTIONS: THE CARBOXYLATION OF D-
RIBULOSE 1,5-BISPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC
CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF
THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH
REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME
ACTIVE SITE.
-1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + CO(2) -> D-3-
PHOSPHO-D-GLYCERATE.
-1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + O(2) ->
PHOSPHO-D-GLYCERATE + 2-PHOSPHOGLYCOLATE.
-1- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS.
-1- SUBCELLULAR LOCATION: CHLOROPLAST.
EMBL: U59822; AAB41157.1;
DR HSSP: P00875; 1RCX
DR MENDEL: 21804; Hilsa.rbcL218
DR PROSITE: PS00157; RUBISCO_LARGE
DR PFAM: PF00016; Rubisco_large
KW Chloroplast; Photosynthesis; Carbon dioxide fixation;
KW Photorespiration; Lyase; Oxidoreductase; Monooxygenase.
FT ACT-SITE 201 201 BINDING OF CO(2) ACTIVATES THE ENZYME

```

FT ACT_SITE 202 202 BINDING OF CO(2) ACTIVATES THE ENZYME.
SQ SEQUENCE 475 AA: 52775 MW: DC4CA579 CR32.

Query Match 45.5% Score 65: DB 8: Length 475:
Best Local Similarity 38.9% Pred. No. 9,08e-00:
Matches 7: Conservative 6: Mismatches 5: Indels 0: Gaps 0:

CU 255 IFARELGVPVNHVDYITG 272
I I I I I I I I I I
CV 1 IFSKNLNKLNMPYIAG 18

RESULT 15
ID 009485 PRELIMINARY: PRT: 475 AA
AC 009485
D: 01-JUL-1997 (FRENBLrel_04, Created)
D: 01-JUL-1997 (FRENBLrel_04, Last sequence update)
DI 01-NOV-1999 (FRENBLrel_12, Last annotation update)
DE RIBULOSE BISPHOSPHATE CARBOXYLASE LARGE CHAIN (EC 4.1.1.39)
DE (FRAGMENT).
GN RBCL
OS Hydrostachys sp. 'Bremer 3089'.
OG Chloroplast.
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
OC euphyllophytes: Spermatophyta: Magnoliophyta: eudicotyledons:
OC core eudicots: Asteridae: Cornales: Hydrostachyaceae: Hydrostachys
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-COLL: BREMER 3089 (UPS)
RA BACKLUND A., BREMER B.
RL Plant Syst. Evol. 200:3-0(1997).
CC -!- FUNCTION: RUBISCO CATALYSES TWO REACTIONS: THE CARBOXYLATION OF D-
RIBULOSE 1,5-BISPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC
CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF
THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH
REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME
ACTIVE SITE.
CC -!- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + CO(2) = 2 3-
PHOSPHO-D-GLYCERATE.
CC -!- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + O(2) = 3-
PHOSPHO-D-GLYCERATE + 2-PROSPHOGLYCOLATE.
CC -!- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS.
CC -!- SUBCELLULAR LOCATION: CHLOROPLAST.
DR EMBL: Y10708; CAA71713.1;
DR MENDEL: 21881; HYDsp:rbcl;21881.
DR PROSITE: PS00157; RUBISCO_LARGE: 1.
DR PFAM: PF00616; RUBISCO_LARGE: 1.
KW Lyase: Chloroplast; Photosynthesis; Carbon dioxide fixation;
KW Photorespiration; Oxidoreductase; Monooxygenase.
FT NON_TER 1
FT ACT_SITE 202 202 BINDING OF CO(2) ACTIVATES THE ENZYME.
SQ SEQUENCE 475 AA: 52775 MW: 4HDC6186 CR32.

Query Match 45.5% Score 65: DB 8: Length 475:
Best Local Similarity 38.9% Pred. No. 9,08e-00:
Matches 7: Conservative 6: Mismatches 5: Indels 0: Gaps 0:

CU 255 IFARELGVPVNHVDYITG 272
I I I I I I I I I I
CV 1 IFSKNLNKLNMPYIAG 18

Search completed: Mon Jun 19 16:13:30 2000
Job time : 11 secs.

1F 10-MAR-1997: 3:0740.
 1R 10-MAR-1996: 2F-080702.
 1A (MEIP) MEJJI MILK PROD CO LTD.
 1I Daitiki K, Iwata A, Kiso K, Kume A, Sone T;
 1R WPI: 97-470495/43.
 1P Peptide immunotherapeutic agent to treat allergic diseases -
 1T contains multi-epitope peptide containing T cell epitope regions
 1I from different allergens
 1R Claim 9: Page 42: 58pp; Japanese.
 1C The present sequence represents a multi-epitope peptide which is used as
 1C a new immunotherapeutic agent. It comprises T cell epitope regions from 2
 1C or more different allergens (preferably linked via arginine or lysine
 1C dimers), where the T cell epitope regions have a positivity index
 1C greater than 1.0 as measured in a patient group responding to the
 1C allergen; have at least 70% reactivity with lymphocytes from patients
 1C responding to the allergen; and are not reactive with immunoglobulin E
 1C (IgE) antibodies from patients responsive to the allergen. The agent can
 1C be used to prevent and treat a wide variety of allergic diseases, e.g. by
 1C desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
 1C Sequence 31 AA;

Query Match: 100.0%; Score 125; DB 1; Length 31;
 Best Local Similarity 100.0%; Pred. No. 1.10e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 SSGKNEGTNIYNNNEAFKVE 20
 ||||| ||||| ||||| |||||
 QY 1 SSGKNEGTNIYNNNEAFKVE 20

RESULT 3
 ID W42121 standard; protein: 354 AA.
 AC W42121:
 SI 15-JUN-1998 (first entry)
 DE Japanese cypress pollen antigen Chaol.
 KW Japanese cypress pollen; antigen; T-cell epitope; Chaol; Chaol2;
 KW diagnosis; allergy; spring tree pollen disease; pollinosis.
 OS Chamaecyparis obtusa.
 RX W09747648-A1.
 1R 18-DEC-1997.
 1R 12-JUN-1997: J:2031.
 1R 14-JUN-1996: JP-133527.
 1A (MEIP) MEJJI MILK PROD CO LTD
 1I Daitiki K, Kiso K.
 1R WPI: 96-052242/05.
 1C T-cell epitope peptide portion of Japanese cypress pollen antigens
 1C Chaol and Chaol2 used for diagnosis and treatment of spring tree
 1C pollen disease
 1C Example 1: Page 14-15: 7pp; Japanese.
 1C The present sequence represents Japanese cypress pollen antigen Chaol.
 1C The present invention describes peptides which correspond to the T-cell
 1C epitope sites on Japanese cypress pollen antigens Chaol and Chaol2. The
 1C peptides can be used as a reagent for the diagnosis of allergy to
 1C Japanese cypress pollen, and as an antigen in the treatment and
 1C prevention of spring tree pollen disease in which the pollinosis
 1C involves reactivity to Japanese cypress pollen.
 1C Sequence 354 AA;

Query Match: 100.0%; Score 125; DB 1; Length 354;
 Best Local Similarity 100.0%; Pred. No. 1.10e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 311 SSGKNEGTNIYNNNEAFKVE 310
 ||||| ||||| ||||| |||||
 QY 1 SSGKNEGTNIYNNNEAFKVE 20

RESULT 4
 ID W04344 standard; protein: 354 AA.
 AC W04344:
 SI 28-NOV-1996 (first entry)
 DE Chamaecyparis obtusa pollen allergen Cha o I (A).
 KW Pollen allergen; Cha o I; T-cell epitope; prevention; treatment;

KW pollinosis.
 OS Chamaecyparis obtusa.
 PN 08176192-A.
 1R 09-JUL-1996.
 1R 21-DEC-1994: 335069.
 1R 21-DEC-1994: JP-335089.
 1A (MEIP) MEJJI MILK PROD CO LTD
 1I WPI: 96-368225/37.
 1R N-PSDB: 138519.
 1C DNA encoding chamaecyparis obtusa pollen allergen - T cell
 1C epitopes) of which are useful in development of preventative and
 1C treating agent for C. obtusa pollen pollinosis
 1C Claim 2: Pages 10-11: 17pp; Japanese.
 1C The present sequence is the C. obtusa pollen allergen Cha o I,
 1C the T-cell epitopes of which can be used in the development of a
 1C preventive and treating agent for C. obtusa pollen pollinosis.
 1C C. obtusa pollen (2.4 kg) was decreased with diethyl ether, and
 1C dried at room temp. overnight. Cha o I was sepd. from it and
 1C purified. RNA was extracted from C. obtusa pollen, and mRNA and
 1C cDNA derived.
 1C Sequence 354 AA;

Query Match: 100.0%; Score 125; DB 1; Length 354;
 Best Local Similarity 100.0%; Pred. No. 1.10e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 311 SSGKNEGTNIYNNNEAFKVE 310
 ||||| ||||| ||||| |||||
 QY 1 SSGKNEGTNIYNNNEAFKVE 20

RESULT 5
 ID W04345 standard; protein: 375 AA.
 AC W04345:
 SI 28-NOV-1996 (first entry)
 DE Chamaecyparis obtusa pollen allergen Cha o I cDNA (B).
 KW Pollen allergen; Cha o I; T-cell epitope; prevention; treatment;
 KW pollinosis.
 OS Chamaecyparis obtusa.
 PN 08176192-A.
 1R 09-JUL-1996.
 1R 21-DEC-1994: 335089.
 1R 21-DEC-1994: JP-335089.
 1A (MEIP) MEJJI MILK PROD CO LTD.
 1I WPI: 96-368225/37.
 1R N-PSDB: 138519.
 1C DNA encoding chamaecyparis obtusa pollen allergen - T cell
 1C epitopes) of which are useful in development of preventative and
 1C treating agent for C. obtusa pollen pollinosis
 1C Claim 8: Pages 11-12: 17pp; Japanese.
 1C The present sequence is the C. obtusa pollen allergen Cha o I,
 1C the T-cell epitopes of which can be used in the development of a
 1C preventive and treating agent for C. obtusa pollen pollinosis.
 1C C. obtusa pollen (2.4 kg) was decreased with diethyl ether, and
 1C dried at room temp. overnight. Cha o I was sepd. from it and
 1C purified. RNA was extracted from C. obtusa pollen, and mRNA and
 1C cDNA derived.
 1C Sequence 375 AA;

Query Match: 100.0%; Score 125; DB 1; Length 375;
 Best Local Similarity 100.0%; Pred. No. 1.10e-06;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 332 SSGKNEGTNIYNNNEAFKVE 351
 ||||| ||||| ||||| |||||
 QY 1 SSGKNEGTNIYNNNEAFKVE 20

RESULT 6
 ID R45577 standard; protein: 367 AA.
 AC R45577:
 SI 13-JUL-1994 (first entry)
 DE Jun 8 11

Query Match 74.48; Score 93; DB 1; Length 20
Best Local Similarity 70.08; Pred. No. 6.61e-03;
Matches 14; Conservative 3; Mismatches 3; Index

PS
Reference Example 6: page 87; 292pp; English.

Query Match 74.48; Score 93; DB 1; Length 20
Best Local Similarity 70.08; Pred. No. 6.61e-03;
Matches 14; Conservative 3; Mismatches 3; Index

CC Synthetic peptides P3-P15 (see W80343-57) are partial sequences of
 CC amino acids 1-157 of the extracellular domain of human Fas, with
 CC between 9 and 11 amino acid residues overlapping one another. P16
 CC (see W80358) is a negative control having no homology with human
 CC Fas. P1-P16 were used in an ELISA, which demonstrated that novel
 CC murine anti-human Fas monoclonal antibody HE7A specifically binds
 CC an amino acid sequence contained in P11. The epitope (see W80350)
 CC was subsequently identified. The invention provides humanised
 CC HE7A antibodies (see W80361-37) produced by CDR grafting. These
 CC antibodies are capable of inducing apoptosis in abnormal cells
 CC expressing Fas, and of inhibiting Fas-induced apoptosis in normal
 CC cells. They are used to evaluate, in animal models, treatments of
 CC diseases that involve Fas/Fas ligand interactions, and also to treat
 CC such diseases.
 CC Sequence 20 AA:

Query Match 74.4% Score 93: DB 1: Length 20:
 Best Local Similarity 70.0% Pred. No. 6.61e-03:
 Matches 14: Conservative 3: Mismatches 3: Indels 0: Gaps 0:

DB 1 SSGKYECCGNIYTKKEAFNVE 20
 ||||| : ||| : ||||| :
 QY 1 SSGKNEGTNIYNNNEAFKVE 20

RESULT 10

ID W80355 standard; peptide: 2: AA:

AC W80355:
 DE 11-JAN-1999 (first entry)
 DE Sugi allergen protein CryJ2 derived epitope for T cells.
 KW T cell epitope; sugi allergen proteins CryJ1; CryJ2: treatment:
 KW sugi-pollinosis; allergic reaction; pollen:
 OS Synthetic.

PN J10259198-A.
 PD 29-SEP-1998.
 PF 22-DEC-1997: 353448.
 PR 24-DEC-1996: JP-34344.
 PA (HAYB.) HAYASHIBARA SEIBUTSU KAGAKU.
 FA (SANY.) SANKYO CO LTD.
 DR WPI: 98-57763/749.
 PI A linked T cell epitope peptide - used for the treatment of
 PI sugi-pollinosis

FS Claim 8: Pages 19-20: 21pp: Japanese.
 CC W80339-58 represent epitopes for T cells, derived from the sugi allergen
 CC proteins CryJ1 (W80339-44, W80350-53 and W80356-58) and CryJ2 (W80345-49
 CC and W80354-55). The peptides are useful for the treatment of
 CC sugi-pollinosis, an allergic reaction of the body to pollen.
 CC Sequence 21 AA:

Query Match 74.4% Score 93: DB 1: Length 21:
 Best Local Similarity 70.0% Pred. No. 6.61e-03:
 Matches 14: Conservative 3: Mismatches 3: Indels 0: Gaps 0:

DB 2 SSGKYECCGNIYTKKEAFNVE 21
 ||||| : ||| : ||||| :
 QY 1 SSGKNEGTNIYNNNEAFKVE 20

RESULT 11

ID R45599 standard; Protein: 26 AA.

AC R45599:
 DE 13-JUL-1994 (first entry)
 DE Cry J 1 pollen allergen peptide CJI-44.1.
 KW Japanese cedar: detection: allergy: treatment: diagnosis:
 KW T cell epitope; sensitivity.

CS Cryptomeria japonica.
 PN W09401560-A.
 PD 20-JAN-1994.
 PF 15-JAN-1993: J00139.
 PR 10-JUL-1992: WC-005661.
 PR 01-SEP-1992: US-938990.
 FA (IMMU-) IMMU-LOGIC PHARM CORP.
 PI Bond JF, Garman RD, Griffith IJ, Kuo M, Pollock J:

DR WPI: 94-035066/04.
 PT Antigens derived from Japanese cedar pollen allergen Cry J 1:
 PT contain at least two T cell epitope(s), used to treat or diagnose
 PT allergy.
 PS Claim 76: Fig 18: 137pp: English.
 CC The sequence is that of an isolated peptide of the Japanese Cedar:
 CC pollen allergen Cry J 1. The peptide, CJI-44.1, can be used for
 CC the treatment and diagnosis of allergies associated with Japanese
 CC cedar pollen. It has enhanced therapeutic properties but reduced
 CC side effects compared to naturally occurring allergens.
 CC Sequence 26 AA:

Query Match 74.4% Score 93: DB 1: Length 26:
 Best Local Similarity 70.0% Pred. No. 6.61e-03:
 Matches 14: Conservative 3: Mismatches 3: Indels 0: Gaps 0:

DB 7 SSGKYECCGNIYTKKEAFNVE 26
 ||||| : ||| : ||||| :
 QY 1 SSGKNEGTNIYNNNEAFKVE 20

RESULT 12

ID R92180 standard; Protein: 28 AA.

AC R92180:
 DE 16-APR-1996 (first entry)
 DE Cry J 1 Japanese Cedar pollen allergen modified peptide (CJI-44.1)
 KW Cry J 1: Japanese cedar pollen allergen; modified: drug product:
 KW allergy: Cryptomeria japonica.
 OS Synthetic.

PN W09527786-A1.
 PD 19-OCT-1995.
 PF 06-APR-1995: U04249.
 PR 08-APR-1994: US-226246.
 PR 06-DEC-1994: US-350225.
 PA (IMMU-) IMMULOGIC PHARM CORP.
 PI Chen X, Evans S, Franzen HM, Kuo M, Powers SP:
 PI Shaked Z:
 DR WPI: 95-366 J1/47.
 PT Modified Cryptomeria japonica (Cry J 1) peptide(s) - useful for
 PT treating allergy to Japanese cedar pollen allergen or
 PT immunologically cross reactive allergens

FS Claim 1: Figure 3: 60pp: English.
 CC Novel peptides of cry J 1 have been modified as a part of a
 CC preformulation scheme to develop an optimised drug product for
 CC therapeutic treatment of humans suffering from allergy to Japanese
 CC cedar pollen allergen or an allergen which is immunologically cross
 CC reactive with Japanese cedar pollen allergen. Such modified peptides
 CC possess certain characteristics which render them particularly
 CC suitable for drug product formulation. Peptide fragments of Cry J 1,
 CC modified and unmodified, are given in W82491-W82525. This peptide
 CC is one of three modified peptides (See W92178-80).

Sequence 28 AA:

Query Match 74.4% Score 93: DB 1: Length 28:
 Best Local Similarity 70.0% Pred. No. 6.61e-03:
 Matches 14: Conservative 3: Mismatches 3: Indels 0: Gaps 0:

DB 9 SSGKYECCGNIYTKKEAFNVE 28
 ||||| : ||| : ||||| :
 QY 1 SSGKNEGTNIYNNNEAFKVE 20

RESULT 13

ID R45598 standard; Protein: 30 AA.

AC R45598:
 DE 13-JUL-1994 (first entry)
 DE Cry J 1 pollen allergen peptide CJI-44.
 KW Japanese cedar: detection: allergy: treatment: diagnosis:
 KW T cell epitope; sensitivity.
 OS Cryptomeria japonica.
 PN W09401560-A.
 PD 20-JAN-1994.
 PF 15-JAN-1993: J00139.

```

PR 10-JUN-1992; WO-005651.
PR 01-SEP-1992; US-938950.
PA (IMM-) IMMUNOLOGIC PHARM CORP.
PI Rond JF, Garman RD, Griffith LG, Kuo M, Pollock J.
DR WPI: 94-035056/24.
PI Antigens derived from Japanese cedar pollen allergen Cryj I -
PI contain at least two T cell epitope(s), used to treat or diagnose
PI allergy.
PS Claim 76; Fig 18; 13pp; English.
CC The sequence is that of an isolated peptide of the Japanese cedar
CC pollen allergen Cryj I. The peptide, CjI-44, can be used for
CC the treatment and diagnosis of allergies associated with Japanese
CC cedar pollen. It has enhanced therapeutic properties but reduced
CC side effects compared to naturally occurring allergens.
SC Sequence 30 AA:

Query Match 74.4% Score 93: DB 1: Length 30:
Best Local Similarity 70.0% Pred. No. 6.61e-03:
Matches 14: Conservative 3: Mismatches 3: Indels 0: Gaps 0:

DB 11 SSGKYEGGNIYIKKEAFNVE 30
QY 1 SSGKNEGTNIYNNNEAFKVE 20

RESULT 14
ID R75388 standard; Protein: 353 AA.
AC R75388:
DE 12-MAY-1996 (first entry)
DE Japanese cedar pollen allergen Cryj I.
KW Japanese cedar pollen allergen; Cryj I; T-cell epitope; peptides;
KW prevention; treatment; cryptomeria pollinosis.
OS Cryptomeria japonica.
PI Key Location/Qualifiers
PI peptide 61..75
F1 peptide /note= "T-cell epitope peptide"
F1 peptide 91..105
F1 peptide /note= "T-cell epitope peptide"
F1 peptide 106..120
F1 peptide /note= "T-cell epitope peptide"
F1 peptide 146..160
F1 peptide /note= "T-cell epitope peptide"
F1 peptide 211..225
F1 peptide /note= "T-cell epitope peptide"
F1 peptide 326..340
F1 peptide /note= "T-cell epitope peptide"
F1 peptide 345..348
F1 peptide /note= "T-cell epitope peptide"
F1 307119295-A.
PN 09-MAY-1995.
PF 20-OCT-1993; 262626.
PR 20-OCT-1993; JP-252526.
PA (MEIP) MEIJI MILK PROD CO LTD.
DR WPI: 95-203834/27.
PI New cryptomeria pollen allergen T-cell epitope peptide - used for
PI prevention, treatment and investigation of Japanese cedar pollinosis
PS Disclosure, Figs 1-2; 8pp; Japanese.
CC R75388 is the Japanese cedar pollen allergen Cryj I, from which the
CC T-cell epitope peptides R89289-R89295 were derived. The peptides
CC can be used for the prevention and treatment of cryptomeria
CC pollinosis, and also for the investigation of pollinosis.
SC Sequence 353 AA:

Query Match 74.4% Score 93: DB 1: Length 353:
Best Local Similarity 70.0% Pred. No. 6.61e-03:
Matches 14: Conservative 3: Mismatches 3: Indels 0: Gaps 0:

DB 311 SSGKYEGGNIYIKKEAFNVE 330
QY 1 SSGKNEGTNIYNNNEAFKVE 20

Search completed: Mon Jun 19 16:15:59 2000
Job time : 6 secs.

```

```

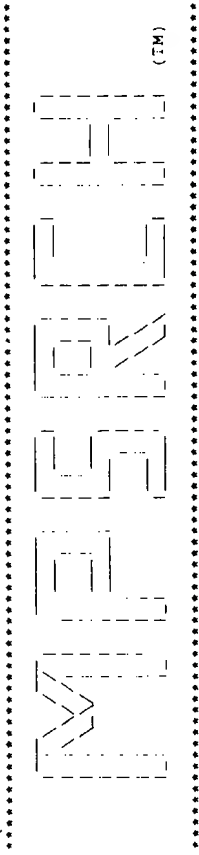
ID R81587 standard; Protein: 353 AA.
AC R81587:
DE 24-MAY-1996 (first entry)
DE Cedar pollen allergen B.
KW Cedar pollen; allergen; immunoglobulin E; IgE; T-cell epitope;
KW antibody; pollinosis; therapy; immunotherapy.
OS Cryptomeria japonica.
PN EP-700929-A2.
PC 13-MAR-1996.
PF 08-SEP-1995; 306295.
PR 10-SEP-1994; JP-242137.
PR 14-JUL-1995; JP-200221.
PR 14-JUL-1995; JP-200204.
PA (HATB) HAYASHIBARA SEIBUTSU KAGAKU.
PI Hino K, Saito S, Taniguchi T.
DR WPI: 95-140976/15.
PI New peptide(s) derived from cedar pollen allergens - activate
PI allergen-specific T-cells, but not allergen-specific IgE antibodies.
PI used for treating cedar pollinosis.
PS Claim 5; Page 31-32; 36pp; English.
CC Synthetic peptides based on portions of cedar pollen allergens A
CC (R81586) and B (R81587) were tested for their ability to activate
CC cedar allergen-specific T-cells but not allergen-specific IgE
CC antibodies. 6 Peptides (R81588-R81595) were identified as T-cell
CC epitopes. These peptides, plus subsequences (R81573-79) essential
CC for T-cell recognition, and homologous peptides (R81588-95) can
CC be used as immunotherapeutic agents to treat or prevent cedar
CC pollinosis, avoiding side-effects such as anaphylaxis.
SC Sequence 353 AA:

Query Match 74.4% Score 93: DB 1: Length 353:
Best Local Similarity 70.0% Pred. No. 6.61e-03:
Matches 14: Conservative 3: Mismatches 3: Indels 0: Gaps 0:

DB 311 SSGKYEGGNIYIKKEAFNVE 330
QY 1 SSGKNEGTNIYNNNEAFKVE 20

Search completed: Mon Jun 19 16:15:59 2000
Job time : 6 secs.

```



Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Msearch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: The Jun 20 13:37:01 2000: MasPar time 2.38 Seconds
Tabular output not generated. 121.365 Million cell updates/sec

Title: >US-09-142-524A-12
Description: (-2C) from US09142524A.pcp
Perfect Score: 125
Sequence: 1 SSGKNEGTNIYNNNEAFKVE 2C

Scoring table: PAM 150
Gap 15

Searched: 145341 seqs, 14437480 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-issued
1:5A_COMB 2:5B_COMB 3:6_COMF 4:PCI_COMB 5:backfiles1

Statistics: Mean 17.862: Variance 55.722: scale 0.321

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

| SUMMARIES | | | | | | | |
|------------|-------|-------------|--------|----|-----------------------------------|-----------|--|
| Result No. | Score | Query Match | Length | DB | Description | Pred. No. | |
| 1 | 59 | 47.2 | 70 | 5 | 5204097-1 Patent No. 5204097 | 1.63e+01 | |
| 2 | 56 | 44.8 | 1365 | 5 | 5194600-4 Patent No. 5194600 | 3.39e+01 | |
| 3 | 55 | 44.0 | 190 | 1 | US-08-849- Sequence 1, Applicati | 4.31e+01 | |
| 4 | 55 | 44.0 | 2710 | 2 | US-08-568- Sequence 12, Applicati | 4.31e+01 | |
| 5 | 55 | 44.0 | 2710 | 2 | US-08-487- Sequence 12, Applicati | 4.31e+01 | |
| 6 | 55 | 44.0 | 3060 | 2 | US-08-487- Sequence 14, Applicati | 4.31e+01 | |
| 7 | 53 | 42.4 | 522 | 1 | US-08-680- Sequence 58, Applicati | 6.95e+01 | |
| 8 | 53 | 42.4 | 1018 | 4 | PCI-US93-1 Sequence 75, Applicati | 6.95e+01 | |
| 9 | 51 | 40.8 | 261 | 3 | US-08-483- Sequence 10, Applicati | 1.11e+02 | |
| 10 | 51 | 40.8 | 1218 | 2 | US-08-400- Sequence 6, Applicati | 1.11e+02 | |
| 11 | 50 | 40.0 | 69 | 2 | US-08-232- Sequence 21, Applicati | 1.41e+02 | |
| 12 | 50 | 40.0 | 69 | 2 | US-08-457- Sequence 21, Applicati | 1.41e+02 | |
| 13 | 50 | 40.0 | 69 | 2 | US-08-467- Sequence 21, Applicati | 1.41e+02 | |
| 14 | 50 | 40.0 | 69 | 3 | US-08-467- Sequence 21, Applicati | 1.41e+02 | |
| 15 | 50 | 40.0 | 321 | 1 | US-08-937- Sequence 3, Applicati | 1.41e+02 | |
| 16 | 50 | 40.0 | 507 | 1 | US-08-305- Sequence 3, Applicati | 1.41e+02 | |
| 17 | 50 | 40.0 | 510 | 2 | US-08-021- Sequence 1, Applicati | 1.41e+02 | |
| 18 | 49 | 39.2 | 293 | 2 | US-08-913- Sequence 2, Applicati | 1.77e+02 | |
| 19 | 49 | 39.2 | 485 | 4 | PCI-US94-0 Sequence 2, Applicati | 1.77e+02 | |
| 20 | 49 | 39.2 | 485 | 3 | US-08-445- Sequence 2, Applicati | 1.77e+02 | |
| 21 | 49 | 39.2 | 527 | 1 | US-08-452- Sequence 10, Applicati | 1.77e+02 | |
| 22 | 49 | 39.2 | 527 | 4 | PCI-US95-0 Sequence 10, Applicati | 1.77e+02 | |
| 23 | 49 | 39.2 | 527 | 1 | US-08-441- Sequence 10, Applicati | 1.77e+02 | |

| | | | | | | | |
|----|----|------|------|---|-----------------------------------|----------|--|
| 24 | 49 | 39.2 | 546 | 2 | US-08-942- Sequence 4, Applicatio | 1.77e+02 | |
| 25 | 49 | 39.2 | 715 | 2 | US-08-849- Sequence 6, Applicatio | 1.77e+02 | |
| 26 | 49 | 39.2 | 1193 | 2 | US-08-400- Sequence 10, Applicati | 1.77e+02 | |
| 27 | 48 | 38.4 | 14 | 5 | 5204097-4 Patent No. 5204097 | 2.21e+02 | |
| 28 | 48 | 38.4 | 20 | 5 | 5204097-3 Patent No. 5204097 | 2.21e+02 | |
| 29 | 48 | 38.4 | 171 | 4 | PCI-US95-0 Sequence 24, Applicati | 2.21e+02 | |
| 30 | 48 | 38.4 | 171 | 2 | US-08-284- Sequence 5, Applicatio | 2.21e+02 | |
| 31 | 48 | 38.4 | 200 | 2 | US-08-709- Sequence 5, Applicatio | 2.21e+02 | |
| 32 | 48 | 38.4 | 232 | 2 | US-08-531- Sequence 5, Applicatio | 2.21e+02 | |
| 33 | 48 | 38.4 | 323 | 3 | US-08-041- Sequence 28, Applicati | 2.21e+02 | |
| 34 | 48 | 38.4 | 336 | 4 | PCI-US92-1 Sequence 3, Applicatio | 2.21e+02 | |
| 35 | 48 | 38.4 | 402 | 2 | US-08-709- Sequence 3, Applicatio | 2.21e+02 | |
| 36 | 48 | 38.4 | 432 | 2 | US-08-813- Sequence 5, Applicatio | 2.21e+02 | |
| 37 | 48 | 38.4 | 432 | 2 | US-08-531- Sequence 2, Applicatio | 2.21e+02 | |
| 38 | 48 | 38.4 | 435 | 1 | US-08-479- Sequence 27, Applicati | 2.21e+02 | |
| 39 | 48 | 38.4 | 435 | 1 | US-08-483- Sequence 27, Applicati | 2.21e+02 | |
| 40 | 48 | 38.4 | 435 | 1 | US-08-361- Sequence 27, Applicati | 2.21e+02 | |
| 41 | 48 | 38.4 | 485 | 4 | PCI-US94-0 Sequence 6, Applicatio | 2.21e+02 | |
| 42 | 48 | 38.4 | 593 | 1 | US-08-478- Sequence 100, Applicat | 2.21e+02 | |
| 43 | 48 | 38.4 | 593 | 1 | US-08-487- Sequence 100, Applicat | 2.21e+02 | |
| 44 | 48 | 38.4 | 593 | 1 | US-08-487- Sequence 4, Applicatio | 2.21e+02 | |
| 45 | 48 | 38.4 | 593 | 2 | US-08-710- Sequence 4, Applicatio | 2.21e+02 | |

AL JNMENIS

RESULT 1
ID 5204097-1 STANDARD: PR 70 AA.
XX xxxxxx
AC
XX
DT
XX
DE Patent No. 5204097
CC
CC Patent No. 5204097
CC APPLICANT: ARNON, RUTH; HARARI, ILANA; KEUSCH, GERALD T.
CC DONOHUE-ROLFE, ARTHUR
CC TITLE OF INVENTION: SHISA TOXIN B CHAIN POLYPEPTIDES AND
CC VACCINE THEREO
CC
CC NUMBER OF SEQUENCES: 5
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/07/464,566
CC FILING DATE: 09-JUN-1989
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 70,247
CC FILING DATE: 06-JUL-1987
CC SEQ ID NO:1
CC LENGTH: 70
CC SEQUENCE 70 AA: 7847 MW: 15614 UN.

| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|----|---------------------------|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|
| CC | CURRENT APPLICATION DATA: | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
|----|---------------------------|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|--|

ASSEMBLY AND USE THEREOF
 NUMBER OF SEQUENCES: 4
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/488,316
 FILING DATE: 05-MAR-1990
 SEQ ID NO:4:
 LENGTH: 1365
 NO SEQUENCE 1365 AA: 156437 MW: 9878636 CN:

Query Match 44.9% Score 55: DB 5: Length 1365:
 Best Local Similarity 46.7% Pred. No. 4,31e+01:
 Matches 7: Conservative 3: Mismatches 5: Indels 0: Gaps 0:

Db 840 QGTNIYNNNEAFKVE 20
 QY 6 EGTNIYNNNEAFKVE 20

RESULT 3
 ID US-08-616-241-1 STANDARD: PRT: 190 AA:

XX xxxxxx

Sequence 1, Application US/08816241

Sequence 1, Application US/08816241

Patent No. 5934,185

GENERAL INFORMATION:

APPLICANT: Bandman, Oida

APPLICANT: Goli, Surya K.

TITLE OF INVENTION: NOVEL RNA EDITING ENZYME

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Inocyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTA for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/816,241

FILING DATE: Filed Herewith

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0239 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 190 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: PROSTUIO9

CLONE: 1645823

NO SEQUENCE 190 AA: 22826 MW: 191242 CN:

Query Match 44.9% Score 55: DB 1: Length 190:
 Best Local Similarity 66.7% Pred. No. 4,31e-01:

Matches 6: Conservative 3: Mismatches 0: Indels 0: Gaps 0:
 Db 161 VYNDNEPK 169
 QY 10 IYNNNEAFK 18

RESULT 4
 ID US-08-568-459A-12 STANDARD: PRT: 2710 AA:

XX xxxxxx

Sequence 12, Application US/08568459A

Sequence 12, Application US/08568459A

Patent No. 5849105

GENERAL INFORMATION:

APPLICANT: Sim, Kim I.

APPLICANT: Chitnis, Chetan

APPLICANT: Miller, Louis R.

APPLICANT: Peterson, David S.

APPLICANT: Sz, Xin-zhaun

APPLICANT: Wellens, Thomas E.

TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX

TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS

NUMBER OF SEQUENCES: 37

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbe Martens Olson & Bear

STREET: 620 Newport Center Drive 19th Floor

CITY: Newport Beach

STATE: California

COUNTRY: US

ZIP: 92650

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #10, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/568,459A

FILING DATE: 07-DEC-1995

CLASSIFICATION: 4A

ATTORNEY/AGENT INFORMATION:

NAME: Knop-Usel, N.J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: 08/010,001

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 235-0550

TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 2710 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHEICAL: NO

ORIGINAL SOURCE:

ORGANISM: Plasmodium falciparum

NO SEQUENCE 2710 AA: 308038 MW: 37145344 CN:

Query Match 44.0% Score 55: DB 2: Length 2710:
 Best Local Similarity 38.5% Pred. No. 4,31e+01:

Matches 5: Conservative 5: Mismatches 3: Indels 0: Gaps 0:

Db 2205 EGKTFIDDDKTEK 2217

QY 6 EGTNIYNNNEAFK 18

RESULT 5

US-08-487-825B-12 STANDARD: PRT: 2710 AA.

Sequence 12, Application US/08487826B

Sequence 12, Application US/08487826B

Patent No. 5993827

GENERAL INFORMATION:

APPLICANT: Sim, Kim L.

APPLICANT: Chitnis, Chetan

APPLICANT: Miller, Louis H.

APPLICANT: Peterson, David S.

APPLICANT: Su, Xin-zhaun

APPLICANT: Welles, Thomas E.

TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX

TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS

NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbe Martens Olson & Bear

STREET: 620 Newport Center Drive 16th Floor

CITY: Newport Beach

STATE: California

COUNTRY: US

ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/487,826B

FILING DATE: 10-SEP-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Israelson, Ned

REGISTRATION NUMBER: 29,655

REFERENCE/DOCKET NUMBER: NIH21.001CPI

TELEPHONE: (619) 235-8550

TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 2710 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORIGINAL SOURCE:

ORGANISM: Plasmodium falciparum

SEQUENCE 2710 AA: 308038 MW: 37145344 CN:

Query Match 44.0% Score 55: DB 2: Length 2710:

Best Local Similarity 38.5%: Pred. No. 4,31e-01:

Matches 5: Conservative 5: Mismatches 3: Indels 0: Gaps 0:

Db 2205 EGKTFDDDKIFK 2217

QY 6 EGTNIYNNNEAFK 18

RESULT 6

ID US-08-487-825B-14 STANDARD: PRT: 3060 AA.

Sequence 14, Application US/08487826B

Sequence 14, Application US/08487826B

Patent No. 5993827

GENERAL INFORMATION:

APPLICANT: Sim, Kim L.

APPLICANT: Chitnis, Chetan

APPLICANT: Miller, Louis H.

APPLICANT: Peterson, David S.

APPLICANT: Su, Xin-zhaun

APPLICANT: Welles, Thomas E.

TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX

TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS

NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbe Martens Olson & Bear

STREET: 620 Newport Center Drive 16th Floor

CITY: Newport Beach

STATE: California

COUNTRY: US

ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/487,826B

FILING DATE: 10-SEP-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Israelson, Ned

REGISTRATION NUMBER: 29,655

REFERENCE/DOCKET NUMBER: NIH21.001CPI

TELEPHONE: (619) 235-8550

TELEFAX: (619) 235-0176

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 3060 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ORIGINAL SOURCE:

ORGANISM: Plasmodium falciparum

SEQUENCE 3060 AA: 4390904 CN:

Query Match 44.0% Score 55: DB 2: Length 3060:

Best Local Similarity 38.5%: Pred. No. 4,31e-01:

Matches 5: Conservative 5: Mismatches 3: Indels 0: Gaps 0:

Db 2205 EGKTFDDDKIFK 2215

QY 6 EGTNIYNNNEAFK 18

RESULT 7

ID US-08-680-726A-58 STANDARD: PRT: 522 AA.

Sequence 58, Application US/08680726A

Sequence 58, Application US/08680726A

Patent No. 5804197

GENERAL INFORMATION:

APPLICANT: Haanes, Elizabeth J.

APPLICANT: Frank, Rexann S.

TITLE OF INVENTION: RECOMBINANT CANINE HERPESVIRUSES

NUMBER OF SEQUENCES: 92

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sheridan Ross & McIntosh

STREET: 1700 Lincoln Street, Suite 3500

CITY: Denver


```

DB      109  DGTSA YTRDGSFGVD 123
          : : : : : : : : : : : :
DB      5  EGTNIYNNNEAFKVE 20

```

```

RESULT 1C
ID US-09-400-159-6 STANDARD; PRT: 1218 AA.
XX
AC XXXXXX
AC

```

Sequence 8, Application: US/08400159
Sequence 6, Application US/08400159
Patent No. 5963282
GENERAL INFORMATION:
APPLICANT: Ish-Horowitz, David
APPLICANT: Henrique, Domingos M.P.
APPLICANT: Lewis, Julian H.
APPLICANT: Myat, Anna M.
APPLICANT: Fleming, Robert J.
APPLICANT: Artavanis-Tsakonas, Spyridon
APPLICANT: Mann, Robert S.
APPLICANT: Gray, Grace E.
TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN
TITLE OF INVENTION: SERPINE GENE AND METHOD
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pencile & Edmonds
STREET: 1355 Avenue of the Americas
CITY: New York
STATE: New York

Query Match 40.8%; Score 51; DB 2; Length 1218;
Best Local Similarity 42.9%; Pred. No. 1.11e+02;
Matches 6; Conservative 4; Mismatches 4; Indels

0b 715 NNGTCYDEGDAFK 728
.:| :|:|:|:
0y 5 NEGNTIYNNNEAFK 18

RESULT 1:
ID JS-C8-292-968-21
XX
AC XXXXX
PRT: 69 AA.

Sequence 21, Application US/08292968

Sequence 21, Application US/08292968

Patent No. 5853122

GENERAL INFORMATION:

APPLICANT: READ, Randy J.

APPLICANT: STEIN, Penelope E.

APPLICANT: COCKLE, Stephen A.

APPLICANT: OOMEN, Raymond P.

APPLICANT: LOOSMORE, Sheena

APPLICANT: KLEIN, Michel H.

APPLICANT: ARMSTRONG, Glen D.

APPLICANT: HAYES, Bart

TITLE OF INVENTION: MODIFICATION OF PER-USSIS TOXIN

NUMBER OF SEQUENCES: 46

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sim & McBurrey

STREET: Suite 701, 330 University Avenue

CITY: Toronto

STATE: Ontario

COUNTRY: Canada

ZIP: M5G 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/292.968

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/110,947

FILING DATE: 24-AUG-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/251,121

FILING DATE: 31-MAY-1993

ATTORNEY/AGENT INFORMATION

NAME: STEWART, Michael

REGISTRATION NUMBER: 2,113

REFERENCE/DOCKET NUMBER: 1038-388

TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 595-1155

TELEFAX: (416) 595-1163

INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:

LENGTH: 69 amino acid

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE 69 AA: 7691 MW: 25533 Da

Query Match 40.0%; Score 50; DB 2; Length 69;
Best Local Similarity 44.4%; Pred. No. 1.41e-02;
Matches 8; Conservative 6; Mismatches 3; Indels

DB 5 TGKVEYTK-YNDDDTFTV 22
::: | | : | : : : | : |
QY 2 SGKNEGTNIYNNNEAFKV 19

RESULT 12
ID US-08-467-974-21 STANDARD; PRT; 69 AA.
XX
AC XXXXXX

Sequence 21, Application US/08467974


```
CC Sequence 21, Application US/08467974
CC Patent No. 5965385
CC GENERAL INFORMATION:
CC APPLICANT: READ, Randy J.
CC APPLICANT: STEIN, Penelope E.
CC APPLICANT: COCKLE, Stephen A.
CC APPLICANT: COHEN, Raymond P.
CC APPLICANT: LOOMORE, Sheena
CC APPLICANT: KLEIN, Michel H.
CC APPLICANT: ARMSTRONG, Glen D.
CC APPLICANT: HAZES, Bart
CC TITLE OF INVENTION: MODIFICATION OF PERTUSSIS TOXIN
CC NUMBER OF SEQUENCES: 46
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Sim & McBurney
CC STREET: Suite 701, 330 University Avenue
CC CITY: Toronto
CC STATE: Ontario
CC COUNTRY: Canada
CC ZIP: M5G 1R7
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/467,974
CC FILING DATE:
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/467,536
CC FILING DATE: 06-JUN-1995
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/292,958
CC FILING DATE: 22-AUG-1994
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/251,121
CC FILING DATE: 31-MAY-1994
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/110,947
CC FILING DATE: 24-AUG-1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: STEWART, Michael I.
CC REGISTRATION NUMBER: 24,973
CC REFERENCE/DOCKET NUMBER: 1038-454 MIS.vg
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (416) 595-1153
CC TELEFAX: (416) 595-1153
CC INFORMATION FOR SEQ ID NO: 21:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 69 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC SEQUENCE 69 AA: 7591 MW: 25533 CN:
SQ
Query Match 40.0% Score 50: DB 2: Length 69:
Best Local Similarity 44.4% Pred. No. 1.41e+02:
Matches 8: Conservative 6: Mismatches 3: Indels 1: Gaps 1:
DB 6 TGRVEYTK-YNDDTFTV 22
QY 2 SGKNEGTNIYNNNEAFKV 19
RESULT 13
ID US-08-467-536-21 STANDARD: PRT: 69 AA.
XX
AC xxxxxx
XX
DT
XX
DE Sequence 21, Application US/08467536
XX Sequence 21, Application US/08467536
XX Patent No. 6018022
```

```
CC Sequence 21, Application US/08467974
CC Patent No. 5977304
CC GENERAL INFORMATION:
CC APPLICANT: READ, Randy J.
CC APPLICANT: STEIN, Penelope E.
CC APPLICANT: COCKLE, Stephen A.
CC APPLICANT: COHEN, Raymond P.
CC APPLICANT: LOOMORE, Sheena
CC APPLICANT: KLEIN, Michel H.
CC APPLICANT: ARMSTRONG, Glen D.
CC APPLICANT: HAZES, Bart
CC TITLE OF INVENTION: MODIFICATION OF PERTUSSIS TOXIN
CC NUMBER OF SEQUENCES: 46
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Sim & McBurney
CC STREET: Suite 701, 330 University Avenue
CC CITY: Toronto
CC STATE: Ontario
CC COUNTRY: Canada
CC ZIP: M5G 1R7
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/467,536
CC FILING DATE: 06-JUN-1995
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/292,958
CC FILING DATE: 22-AUG-1994
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/251,121
CC FILING DATE: 31-MAY-1994
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/110,947
CC FILING DATE: 24-AUG-1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: STEWART, Michael I.
CC REGISTRATION NUMBER: 24,973
CC REFERENCE/DOCKET NUMBER: 1038-454 MIS.vg
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (416) 595-1153
CC TELEFAX: (416) 595-1153
CC INFORMATION FOR SEQ ID NO: 21:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 69 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC SEQUENCE 69 AA: 7591 MW: 25533 CN:
SQ
Query Match 40.0% Score 50: DB 2: Length 69:
Best Local Similarity 44.4% Pred. No. 1.41e+02:
Matches 8: Conservative 6: Mismatches 3: Indels 1: Gaps 1:
DB 6 TGRVEYTK-YNDDTFTV 22
QY 2 SGKNEGTNIYNNNEAFKV 19
RESULT 14
ID US-08-467-976-21 STANDARD: PRT: 69 AA.
XX
AC xxxxxx
XX
DT
XX
DE Sequence 21, Application US/08467976
XX Sequence 21, Application US/08467976
XX Patent No. 6018022
```

CC GENERAL INFORMATION:
CC APPLICANT: READ, Randy J.
CC APPLICANT: STEIN, Penelope E.
CC APPLICANT: COCKLE, Stephen A.
CC APPLICANT: COHEN, Raymond P.
CC APPLICANT: LOOSMORE, Sheena
CC APPLICANT: KLEIN, Michael H.
CC APPLICANT: ARMSTRONG, Glen D.
CC APPLICANT: HAZES, Bart
CC TITLE OF INVENTION: MODIFICATION OF PERTUSSIS TOXIN
CC NUMBER OF SEQUENCES: 46
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Sir & McBurney
CC STREET: Suite 701, 330 University Avenue
CC City: Toronto
CC STATE: Ontario
CC COUNTRY: Canada
CC ZIP: M5G 1K7
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/457,976
CC FILING DATE: 06-JUN-1995
CC CLASSIFICATION: 435
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/292,968
CC FILING DATE: 22-AUG-1994
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/251,121
CC FILING DATE: 31-MAY-1994
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 08/110,947
CC FILING DATE: 24-AUG-1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: STEWART, Michael I.
CC REGISTRATION NUMBER: 24,973
CC REFERENCE/DOCKET NUMBER: 1038-453 M 81vg
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (416) 595-1155
CC TELEFAX: (416) 595-1163
CC INFORMATION FOR SEQ ID NO: 21:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 69 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC SEQUENCE 69 AA: 7691 MW: 25533 CN:

Db 6 TGKVEYK-VNDDTFTV 22
QY 2 SKNGEINYNNEAFKV 19

RESULT 15
IL US-08-937-972-3 STANDARD: PRT: 321 AA.
XX
AC xxxxxx
XX
DE
XX
DE
XX
XX
Sequence 3, Application US/08937972
XX
XX
Sequence 3, Application US/08937972
XX
XX
Patent No. 5932443
XX
XX
GENERAL INFORMATION:
XX APPLICANT: Lal, Preeti

CC APPLICANT: Bandman, Olga
CC APPLICANT: Corley, Neil C.
CC APPLICANT: Shah, Purvi
CC TITLE OF INVENTION: ANTIGENS
CC NUMBER OF SEQUENCES: 6
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Incyte Pharmaceuticals, Inc.
CC STREET: 3174 Porter Drive
CC City: Palo Alto
CC STATE: CA
CC COUNTRY: USA
CC ZIP: 94304
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Diskette
CC COMPUTER: IBM Compatib.e
CC OPERATING SYSTEM: DOS
CC SOFTWARE: FastSeq for Windows Version 2.0
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/937,972
CC FILING DATE: Herewith
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER:
CC FILING DATE:
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Billings, Lucy J.
CC REGISTRATION NUMBER: 35,749
CC REFERENCE/DOCKET NUMBER: PF-0400 US
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 650-855-0555
CC TELEFAX: 650-845-4156
CC TELEX:
CC INFORMATION FOR SEQ ID NO: 3:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 321 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC IMMEDIATE SOURCE:
CC LIBRARY: BLADNOT04
CC CLONE: 13:8190
CC SEQUENCE 321 AA: 35892 MW: 544482 PN:

Query Match 40.0% Score 691 IP: 40 Length 321
Best Local Similarity 30.8% Prod No: 140012
Matches 4: Conservation 40 Mismatches 3: Indels 0
Db 171 INVEGPHENFNE 183
QY 8 TNINYNNEAFKE 20

Search completed: Tue Jun 20 13:37:07 2000
Job time : 6 secs.

[REDACTED]

(TM)

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MFsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Jun 19 16:27:44 2000; MasPar time 15.63 Seconds
Total output not generated. 128,946 Million cell updates/sec

Title: <US-09-142-524A-12
Description: (1-20) from US09142524A.pap
Perfect Score: 125
Sequence: 1 SSGKNEGTNIYNNNEAFKVE 20

Scoring table: PAM 150
Gap 15

Searched: 721208 seqs, 100765575 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-pending
1: PCT 2:06 3:056 4:07 5:080 6:081 7:082 8:083 9:084
10:084B 11:085 12:086 13:087 14:088 15:089 16:090 17:091
18:092 19:093 20:094 21:095 22:NEWP 23:NEWU60 24:NEWU8
25:NEWU9

Statistics: Mean 20.906; Variance 54.342; scale 0.385

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

| SUMMARIES | | | | |
|------------|-------|-------------|--------|--|
| Result No. | Score | Query Match | Length | DR ID Description Pred. No. |
| 1 | 125 | 100.0 | 31 | US-09-142- Sequence 5, Applicatio 1.39e-07 |
| 2 | 108 | 96.4 | 346 | US-60-109- Sequence 1, Applicatio 3.54e-05 |
| 3 | 108 | 96.4 | 367 | US-08-226- Sequence 95, Applicatio 3.54e-05 |
| 4 | 108 | 86.4 | 367 | US-08-467- Sequence 95, Applicatio 3.54e-05 |
| 5 | 108 | 86.4 | 367 | US-08-457- Sequence 95, Applicatio 3.54e-05 |
| 6 | 108 | 86.4 | 367 | US-08-467- Sequence 95, Applicatio 3.54e-05 |
| 7 | 108 | 86.4 | 367 | US-08-350- Sequence 95, Applicatio 3.54e-05 |
| 8 | 108 | 86.4 | 367 | US-08-468- Sequence 95, Applicatio 3.54e-05 |
| 9 | 93 | 74.4 | 20 | US-07-938- Sequence 57, Applicatio 4.09e-03 |
| 10 | 93 | 74.4 | 20 | US-08-468- Sequence 57, Applicatio 4.09e-03 |
| 11 | 93 | 74.4 | 20 | US-08-350- Sequence 57, Applicatio 4.09e-03 |
| 12 | 93 | 74.4 | 20 | US-08-226- Sequence 57, Applicatio 4.09e-03 |
| 13 | 93 | 74.4 | 26 | US-08-467- Sequence 91, Applicatio 4.09e-03 |
| 14 | 93 | 74.4 | 26 | US-08-468- Sequence 91, Applicatio 4.09e-03 |
| 15 | 93 | 74.4 | 26 | US-08-350- Sequence 91, Applicatio 4.09e-03 |
| 16 | 93 | 74.4 | 26 | US-08-467- Sequence 132, Applicatio 4.09e-03 |
| 17 | 93 | 74.4 | 28 | US-08-467- Sequence 132, Applicatio 4.09e-03 |
| 18 | 93 | 74.4 | 28 | US-08-467- Sequence 131, Applicatio 4.09e-03 |
| 19 | 93 | 74.4 | 29 | US-08-467- Sequence 131, Applicatio 4.09e-03 |
| 20 | 93 | 74.4 | 29 | US-08-458- Sequence 131, Applicatio 4.09e-03 |

| | | | | | |
|----|----|------|-----|----|--|
| 21 | 93 | 74.4 | 30 | 10 | US-08-457- Sequence 261, Applicat 4.09e-03 |
| 22 | 93 | 74.4 | 30 | 10 | US-08-457- Sequence 261, Applicat 4.09e-03 |
| 23 | 93 | 74.4 | 30 | 10 | US-08-457- Sequence 261, Applicat 4.09e-03 |
| 24 | 93 | 74.4 | 30 | 10 | US-08-457- Sequence 260, Applicat 4.09e-03 |
| 25 | 93 | 74.4 | 30 | 10 | US-08-457- Sequence 260, Applicat 4.09e-03 |
| 26 | 93 | 74.4 | 30 | 10 | US-08-457- Sequence 260, Applicat 4.09e-03 |
| 27 | 93 | 74.4 | 30 | 8 | US-08-350- Sequence 260, Applicat 4.09e-03 |
| 28 | 93 | 74.4 | 30 | 10 | US-08-458- Sequence 90, Applicati 4.09e-03 |
| 29 | 93 | 74.4 | 40 | 8 | US-08-350- Sequence 70, Applicati 4.09e-03 |
| 30 | 93 | 74.4 | 40 | 13 | US-08-459- Sequence 70, Applicati 4.09e-03 |
| 31 | 93 | 74.4 | 40 | 13 | US-08-457- Sequence 70, Applicati 4.09e-03 |
| 32 | 93 | 74.4 | 40 | 13 | US-08-457- Sequence 70, Applicati 4.09e-03 |
| 33 | 93 | 74.4 | 63 | 10 | US-08-458- Sequence 65, Applicati 4.09e-03 |
| 34 | 93 | 74.4 | 63 | 10 | US-08-457- Sequence 65, Applicati 4.09e-03 |
| 35 | 93 | 74.4 | 63 | 10 | US-08-457- Sequence 65, Applicati 4.09e-03 |
| 36 | 93 | 74.4 | 63 | 10 | US-08-457- Sequence 65, Applicati 4.09e-03 |
| 37 | 93 | 74.4 | 63 | 7 | US-08-226- Sequence 65, Applicati 4.09e-03 |
| 38 | 93 | 74.4 | 374 | 10 | US-08-457- Sequence 2, Applicatio 4.09e-03 |
| 39 | 93 | 74.4 | 374 | 10 | US-08-457- Sequence 2, Applicatio 4.09e-03 |
| 40 | 93 | 74.4 | 374 | 10 | US-08-457- Sequence 2, Applicatio 4.09e-03 |
| 41 | 93 | 74.4 | 374 | 7 | US-08-226- Sequence 2, Applicatio 4.09e-03 |
| 42 | 93 | 74.4 | 374 | 8 | US-08-350- Sequence 2, Applicatio 4.09e-03 |
| 43 | 93 | 74.4 | 374 | 4 | US-07-730- Sequence 2, Applicatio 4.09e-03 |
| 44 | 93 | 74.4 | 374 | 4 | US-07-729- Sequence 2, Applicatio 4.09e-03 |
| 45 | 93 | 74.4 | 374 | 10 | US-08-458- Sequence 2, Applicatio 4.09e-03 |

ALIGNMENTS

RESULT 1
ID US-09-142-524-5 STANDARD: PRT: 31 AA.
XX
AC xxxxxx
DT
DX
XX
Sequence 5, Application US/ 142544
Sequence 5, Application US/09.42024
GENERAL INFORMATION:
APPLICANT: Sone, Toshio
APPLICANT: Kume, Akio
APPLICANT: Kairiki, Kazuo
APPLICANT: Iwama, Akio
APPLICANT: Kiro, Kohsuke
TITLE OF INVENTION: Peptide-based immunomodulatory Agent for treating
FILE REFERENCE: Docket No. 99-0133
CURRENT APPLICATION NUMBER: US/ 97/42 124
EARLIER FILING DATE: 1999-01-04
EARLIER APPLICATION NUMBER: 8/63/762
EARLIER FILING DATE: 1996-03-10
EARLIER APPLICATION NUMBER: PCT/JF97/00740
NUMBER OF SEQ ID NOS: 5
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 31
TYPE: PRT
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Description of Unknown Organism:peptide
SEQUENCE 31 AA: 3585 MW: 4902 CN;

Query Match 100.0%; Score 125; DB 17; Length 31;
Best Local Similarity 100.0%; Pred. No. 1.39e-07;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 SSGKNEGTNIYNNNEAFKVE 20
IIIIIIIIII IIIIIIIII
QY 1 SSGKNEGTNIYNNNEAFKVE 20

CC FILING DATE: December 6, 1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Jane E. Remillard
CC REGISTRATION NUMBER: 38,872
CC REFERENCE/DOCKET NUMBER: 025.6 USD2 (IMI-028CPD2)
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 227-7400
CC TELEFAX: (617) 227-5941
CC INFORMATION FOR SEQ ID NO: 95:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 367 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 367 AA: 39846 MW: 645078 CN:
SC
Query Match 86.4% Score 108 DB 10 Length 367
Best Local Similarity 85.0% Pred. No. 3 54e-05
Matches 17: Conservative 1: Mismatches 2: Indels 0: Gaps 0:
DB 332 SSGKTEETNIYNSNEAFKVE 351
QY 1 SSGKNEGNIYNNNEAFKVE 20
RESULT 5
ID US-08-467-006-95 STANDARD PRT: 367 AA.
XX
AC xxxxxx
DT
DE
XX
Sequence 95, Application US/08467006
Sequence 95, Application US/08467006
GENERAL INFORMATION:
CC APPLICANT: Griffith, Irwin J.
CC APPLICANT: Pollock, Joanne
CC APPLICANT: Bond, Julian F.
CC APPLICANT: Garman, Richard D.
CC APPLICANT: Kuo, Mei-Chang
CC APPLICANT: Yeung, Siu-mei H.
CC APPLICANT: Brauer, Andrew
CC APPLICANT: Exley, Mark A.
CC APPLICANT: Powers, Steven P.
CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
CC TITLE OF INVENTION: Japanese Cedar Pollen
CC NUMBER OF SEQUENCES: 261
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
CC STREET: 610 Lincoln St
CC CITY: Waltham
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02154
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/467.006
CC FILING DATE: June 6, 1995
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/350,225
CC FILING DATE: December 6, 1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Jane E. Remillard
CC REGISTRATION NUMBER: 38,872
CC REFERENCE/DOCKET NUMBER: 025.6 USD6 (IMI-028CPD6)
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 227-7400

CC TELEFAX: (617) 227-5941
CC INFORMATION FOR SEQ ID NO: 95:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 367 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 367 AA: 39846 MW: 645078 CN:
SC
Query Match 86.4% Score 108 DB 10 Length 367
Best Local Similarity 85.0% Pred. No. 3 54e-05
Matches 17: Conservative 1: Mismatches 2: Indels 0: Gaps 0:
DB 332 SSGKTEETNIYNSNEAFKVE 351
QY 1 SSGKNEGNIYNNNEAFKVE 20
RESULT 6
ID US-08-467-697-95 STANDARD PRT: 367 AA.
XX
AC xxxxxx
DT
DE
XX
Sequence 95, Application US/08467697
Sequence 95, Application US/08467697
GENERAL INFORMATION:
CC APPLICANT: Griffith, Irwin J.
CC APPLICANT: Pollock, Joanne
CC APPLICANT: Bond, Julian F.
CC APPLICANT: Garman, Richard D.
CC APPLICANT: Kuo, Mei-Chang
CC APPLICANT: Yeung, Siu-mei H.
CC APPLICANT: Brauer, Andrew
CC APPLICANT: Exley, Mark A.
CC APPLICANT: Powers, Steven P.
CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
CC TITLE OF INVENTION: Japanese Cedar Pollen
CC NUMBER OF SEQUENCES: 261
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
CC STREET: 610 Lincoln St
CC CITY: Waltham
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02154
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/467.697
CC FILING DATE: June 6, 1995
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/350,225
CC FILING DATE: December 6, 1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Jane E. Remillard
CC REGISTRATION NUMBER: 38,872
CC REFERENCE/DOCKET NUMBER: 025.6 USD4 (IMI-028CPD4)
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 227-7400
CC TELEFAX: (617) 227-5941
CC INFORMATION FOR SEQ ID NO: 95:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 367 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein

Matches 17: Conservative 1: Mismatch 2: Indels 0: Gaps 0:

Db 332 SSGKTEINYNNEAFKVE 35:
||||| ||||| |||||
QY 1 SSGKNEGNIYNNNEAFKVE 20

RESULT 9
ID US-08-467-947-57 STANDARD: PRT: 20 AA.

XX xxxxxx

Sequence 57, Application US/08457697

Sequence 57, Application US/08457697

GENERAL INFORMATION:
APPLICANT: Griffith, Irwin J.
APPLICANT: Pollock, Joanne
APPLICANT: Bond, Julian F.
APPLICANT: Garman, Richard D.
APPLICANT: Kuo, Mei-Chang
APPLICANT: Yeung, Siu-mei H.
APPLICANT: Breder, Andrew
APPLICANT: Exley, Mark A.
APPLICANT: Powers, Steven P.
TITLE OF INVENTION: Allergenic Proteins And Peptides From
TITLE OF INVENTION: Japanese Cedar Pollen
NUMBER OF SEQUENCES: 261
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
STREET: 610 Lincoln St
CITY: Waltham
STATE: MA
COUNTRY: USA
ZIP: 02154

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,697
FILING DATE: June 6, 1995

CLASSIFICATION: 424

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/350,275
FILING DATE: December 6, 1994

ATTORNEY/AGENT INFORMATION:
NAME: Jane E. Remillard
REGISTRATION NUMBER: 38,872
REFERENCE/DOCKET NUMBER: 025.6 USD4 (IM-028CPD4)

TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 57:

SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid

MOLECULE TYPE: peptide
FRAGMENT TYPE: internal

SEQUENCE 20 AA: 2221 MW: 2174 CN:

Query Match 74.4%: Score 93: DB 10: Length 20:

Best Local Similarity 70.0%: Pred. No. 4.09e-03:

Matches 14: Conservative 3: Mismatches 3: Indels 0: Gaps 0:

Db 1 SSGKTEINYNNEAFKVE 20

QY 1 SSGKNEGNIYNNNEAFKVE 20

RESULT 11

ID US-08-468-940-57 STANDARD: PRT: 20 AA.

XX xxxxxx

AC xxxxxx

DT ||||| ||||| |||||

XX 1 SSGKNEGNIYNNNEAFKVE 20

QY 1 SSGKNEGNIYNNNEAFKVE 20

RESULT 10

ID US-07-938-990A-57 STANDARD: PRT: 20 AA.

AC xxxxxx

XX xxxxxx

DT ||||| ||||| |||||

XX Sequence 57, Application US/07938-990A

Sequence 57, Application US/07938-990A

GENERAL INFORMATION:

APPLICANT: Griffith, Irwin J.

APPLICANT: Pollock, Joanne

APPLICANT: Bond, Julian

TITLE OF INVENTION: Allergenic Proteins And Peptides From

TITLE OF INVENTION: Japanese Cedar Pollen

NUMBER OF SEQUENCES: 79

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lahive & Cockfield

STREET: Sixty State Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/938-990A

FILING DATE: 19920901

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/000,134

FILING DATE: July 15, 1991

ATTORNEY/AGENT INFORMATION:

NAME: Amy E. Mandragoras

REGISTRATION NUMBER: 35,237

REFERENCE/DOCKET NUMBER: IPC-0250C (IM-028)

TELEPHONE: (617) 227-7400

TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 10

SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: peptide

FRAGMENT TYPE: internal

SEQUENCE 20 AA: 2221 MW: 2174 CN:

Query Match 74.4%: Score 93: DB 4: Length 20:

Best Local Similarity 70.0%: Pred. No. 4.09e-03:

Matches 14: Conservative 3: Mismatches 3: Indels 0: Gaps 0:

Db 1 SSGKTEGNIYKKEAFNVE 20

QY 1 SSGKNEGNIYNNNEAFKVE 20

RESULT 11

ID US-08-468-940-57 STANDARD: PRT: 20 AA.

XX xxxxxx

AC xxxxxx

DT ||||| ||||| |||||

XX Sequence 57, Application US/08468940

| | | | |
|---|---|--|---|
| XX | Sequence 57, Application US/0846894C | XX | Sequence 57, Application US/08350225 |
| CC | GENERAL INFORMATION: | CC | GENERAL INFORMATION: |
| CC | APPLICANT: Griffith, Irwin J. | CC | APPLICANT: Griffith, Irwin J. |
| CC | APPLICANT: Pollock, Joanne | CC | APPLICANT: Pollock, Joanne |
| CC | APPLICANT: Bond, Julian F. | CC | APPLICANT: Bond, Julian F. |
| CC | APPLICANT: Garmag, Richard D. | CC | APPLICANT: Garmag, Richard D. |
| CC | APPLICANT: Kuo, Wei-Chang | CC | APPLICANT: Kuo, Wei-Chang |
| CC | APPLICANT: Yeung, Siu-mei H. | CC | APPLICANT: Yeung, Siu-mei H. |
| CC | APPLICANT: Brauer, Andrew | CC | APPLICANT: Brauer, Andrew |
| CC | APPLICANT: Exley, Mark A. | CC | APPLICANT: Exley, Mark A. |
| CC | APPLICANT: Powers, Steven P. | CC | APPLICANT: Powers, Steven P. |
| CC | TITLE OF INVENTION: Allergenic Proteins And Peptides From | CC | TITLE OF INVENTION: Allergenic Proteins And Peptides From |
| CC | TITLE OF INVENTION: Japanese Cedar Pollen | CC | TITLE OF INVENTION: Japanese Cedar Pollen |
| CC | NUMBER OF SEQUENCES: 261 | CC | NUMBER OF SEQUENCES: 261 |
| CC | CORRESPONDENCE ADDRESS: | CC | CORRESPONDENCE ADDRESS: |
| CC | ADDRESSEE: Immologic Pharmaceutical Corporation, Inc. | CC | ADDRESSEE: Immologic Pharmaceutical Corporation, Inc. |
| CC | STREET: 610 Lincoln St | CC | STREET: 610 Lincoln St |
| CC | CITY: Waltham | CC | CITY: Waltham |
| CC | STATE: MA | CC | STATE: MA |
| CC | COUNTRY: USA | CC | COUNTRY: USA |
| CC | ZIP: 02154 | CC | ZIP: 02154 |
| CC | COMPUTER READABLE FORM: | CC | COMPUTER READABLE FORM: |
| CC | MEDIUM TYPE: Floppy disk | CC | MEDIUM TYPE: Floppy disk |
| CC | COMPUTER: IBM PC compatible | CC | COMPUTER: IBM PC compatible |
| CC | OPERATING SYSTEM: PC-DOS/MS-DOS | CC | OPERATING SYSTEM: PC-DOS/MS-DOS |
| CC | SOFTWARE: Patent In Release #1.0. Version #1.25 | CC | SOFTWARE: Patent In Release #1.0. Version #1.25 |
| CC | CURRENT APPLICATION DATA: | CC | CURRENT APPLICATION DATA: |
| CC | APPLICATION NUMBER: US/08468,940 | CC | APPLICATION NUMBER: US/08468,940 |
| CC | FILING DATE: 6, 1994 | CC | FILING DATE: 6, 1994 |
| CC | CLASSIFICATION: 424 | CC | CLASSIFICATION: 424 |
| CC | PRIOR APPLICATION DATA: | CC | PRIOR APPLICATION DATA: |
| CC | APPLICATION NUMBER: 08/450,225 | CC | APPLICATION NUMBER: 08/450,225 |
| CC | FILING DATE: December 6, 1994 | CC | FILING DATE: December 6, 1994 |
| CC | APPLICATION NUMBER: 08/226,248 | CC | APPLICATION NUMBER: 08/226,248 |
| CC | FILING DATE: April 8, 1994 | CC | FILING DATE: April 8, 1994 |
| CC | APPLICATION NUMBER: 07/338,990 | CC | APPLICATION NUMBER: 07/338,990 |
| CC | FILING DATE: September 1, 1992 | CC | FILING DATE: September 1, 1992 |
| CC | APPLICATION NUMBER: PCT/US93/00139 | CC | APPLICATION NUMBER: PCT/US93/00139 |
| CC | FILING DATE: January 15, 1993 | CC | FILING DATE: January 15, 1993 |
| CC | ATTORNEY/AGENT INFORMATION: | CC | ATTORNEY/AGENT INFORMATION: |
| CC | NAME: Darlene A. Vanstone | CC | NAME: Darlene A. Vanstone |
| CC | REGISTRATION NUMBER: 35,729 | CC | REGISTRATION NUMBER: 35,729 |
| CC | REFERENCE/DOCKET NUMBER: 035.6 US (IM-028CP2) | CC | REFERENCE/DOCKET NUMBER: 035.6 US (IM-028CP2) |
| CC | TELECOMMUNICATION INFORMATION: | CC | TELECOMMUNICATION INFORMATION: |
| CC | TELEPHONE: (617) 466-6060 | CC | TELEPHONE: (617) 466-6060 |
| CC | TELEFAX: (617) 466-6040 | CC | TELEFAX: (617) 466-6040 |
| CC | INFORMATION FOR SEQ ID NO: 1 | CC | INFORMATION FOR SEQ ID NO: 1 |
| CC | SEQUENCE CHARACTERISTICS: | CC | SEQUENCE CHARACTERISTICS: |
| CC | LENGTH: 20 amino acids | CC | LENGTH: 20 amino acids |
| CC | TYPE: amino acid | CC | TYPE: amino acid |
| CC | TOPOLOGY: linear | CC | TOPOLOGY: linear |
| CC | MOLECULE TYPE: Peptide | CC | MOLECULE TYPE: Peptide |
| CC | FRAGMENT TYPE: Internal | CC | FRAGMENT TYPE: Internal |
| CC | SEQUENCE 20 AA: 2221 MW: 2174 CN: | CC | SEQUENCE 20 AA: 2221 MW: 2174 CN: |
| CC | Query Match 74.4% Score 93: DB 10: Length 20: | CC | Query Match 74.4% Score 93: DB 8: Length 20: |
| CC | Best Local Similarity 70.0% Pred. No. 4.09e-03: | CC | Best Local Similarity 70.0% Pred. No. 4.09e-03: |
| CC | Matches 14: Conservative 3: Mismatches 3: Indels 0: | CC | Matches 14: Conservative 3: Mismatches 3: Indels 0: |
| DB | 1 SSGKEGCGNTYTKKEAFNVE 20 | DB | 1 SSGKYESGNIYTKKEAFNVE 20 |
| OY | 1 SSGKNEGTNIYNNNEAFKVE 20 | OY | 1 SSGKNEGTNIYNNNEAFKVE 20 |
| RESULT 12 | | RESULT 13 | |
| ID US-08-350-225-57 STANDARD: PRT: 20 AA. | | ID US-08-226-248A-57 STANDARD: PRT: 20 AA. | |
| XX | xxxxxx | XX | xxxxxx |
| AC | | AC | |
| XX | | XX | |
| DT | | DT | |
| XX | | XX | |
| DE | Sequence 57, Application US/08350225 | DE | Sequence 57, Application US/08226248A |
| CC | | CC | |


```

CC GENERAL INFORMATION:
CC APPLICANT: Griffith, Irwin J.:
CC APPLICANT: Pollock, Joanne:
CC APPLICANT: Bond, Julian F.:
CC APPLICANT: Garman, Richard D.:
CC APPLICANT: Kuo, Mei-Chang:
CC APPLICANT: Yeung, Siu-mei H.:
CC APPLICANT: Brauer, Andrew:
CC APPLICANT: Exley, Mark A.:
CC APPLICANT: Powers, Steven P.:
CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
CC TITLE OF INVENTION: Japanese Cedar Pollen
CC NUMBER OF SEQUENCES: 201
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
CC STREET: 610 Lincoln St
CC CITY: Waltham
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02154
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/09/226,248A
CC FILING DATE: April 8, 1994
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/938,990
CC FILING DATE: September 1, 1992
CC APPLICATION NUMBER: PCT/US93/00139
CC FILING DATE: January 15, 1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Darlene A. Vanstone
CC REGISTRATION NUMBER: 35,729
CC REFERENCE/DOCKET NUMBER: 025.5 US (IM1-028CP)
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 466-6000
CC TELEFAX: (617) 466-6040
CC INFORMATION FOR SEQ ID NO: 57:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 20 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC FRAGMENT TYPE: internal
CC SEQUENCE: 20 AA: 2221 MW: 2174 CN:

Db Query Match 74.4% Score 93: DB 7: Length 20:
Best Local Similarity 70.0% Pred. No. 4: 09e-03:
Matches 14: Conservative 3: Mismatches 3: Indels 0: Gaps 0:

QY 1 SSGKVEGGNIYTKKEAFNVE 20
|||||i|||:||||:|
1 SSGKNEGTNIYNNNEAFKVE 20

RESULT 14
ID US-08-226-248A-91 STANDARD: PRT: 26 AA.
XX
AC xxxxxx
DE
DT
XX
XX
Sequence 91, Application US/08226248A
DE
DE
Sequence 91, Application US/08226248A
CC GENERAL INFORMATION:
CC APPLICANT: Griffith, Irwin J.:
CC APPLICANT: Pollock, Joanne:
CC APPLICANT: Bond, Julian F.:

```

```

CC APPLICANT: Garman, Richard D.:
CC APPLICANT: Kuo, Mei-Chang:
CC APPLICANT: Yeung, Siu-mei H.:
CC APPLICANT: Brauer, Andrew:
CC APPLICANT: Exley, Mark A.:
CC APPLICANT: Powers, Steven P.:
CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
CC TITLE OF INVENTION: Japanese Cedar Pollen
CC NUMBER OF SEQUENCES: 201
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
CC STREET: 610 Lincoln St
CC CITY: Waltham
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02154
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/09/226,248A
CC FILING DATE: April 8, 1994
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/938,990
CC FILING DATE: September 1, 1992
CC APPLICATION NUMBER: PCT/US93/00139
CC FILING DATE: January 15, 1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Darlene A. Vanstone
CC REGISTRATION NUMBER: 35,729
CC REFERENCE/DOCKET NUMBER: 025.5 US (IM1-028CP)
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 466-6000
CC TELEFAX: (617) 466-6040
CC INFORMATION FOR SEQ ID NO: 91:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 26 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC FRAGMENT TYPE: internal
CC SEQUENCE: 26 AA: 2873 MW: 3771 CN:

Db Query Match 74.4% Score 93: DB 7: Length 26:
Best Local Similarity 70.0% Pred. No. 4: 09e-03:
Matches 14: Conservative 3: Mismatches 3: Indels 0: Gaps 0:

QY 1 SSGKNEGTNIYNNNEAFKVE 20
|||||i|||:||||:|
1 SSGKVEGGNIYTKKEAFNVE 20

RESULT 15
ID US-08-467-697-91 STANDARD: PRT: 26 AA.
XX
AC xxxxxx
DE
DT
XX
XX
Sequence 91, Application US/08467697
DE
DE
Sequence 91, Application US/08467697
CC GENERAL INFORMATION:
CC APPLICANT: Griffith, Irwin J.:
CC APPLICANT: Pollock, Joanne:
CC APPLICANT: Bond, Julian F.:
CC APPLICANT: Garman, Richard D.:
CC APPLICANT: Kuo, Mei-Chang:
CC APPLICANT: Yeung, Siu-mei H.:
CC APPLICANT: Brauer, Andrew:

```

APPLICANT: Exley, Mark A.
 APPLICANT: Powers, Steven P.
 TITLE OF INVENTION: Allergenic Proteins And Peptides From
 TITLE OF INVENTION: Japanese Cedar Pollen
 NUMBER OF SEQUENCES: 261
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
 STREET: 610 Lincoln St
 CITY: Waltham
 STATE: MA
 COUNTRY: USA
 ZIP: 02154
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent-In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/467,697
 FILING DATE: June 6, 1995
 CLASSIFICATION: 424
 PRICE APPLICATION DATA:
 APPLICATION NUMBER: 08/350,225
 FILING DATE: December 6, 1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Jane E. Remillard
 REGISTRATION NUMBER: 38,872
 REFERENCE/DOCKET NUMBER: 025.6 USDA (IM1-028CPD4)
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 227-7400
 TELEFAX: (617) 227-5941
 INFORMATION FOR SEQ ID NO: 91:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 26 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 FRAGMENT TYPE: internal
 SEQUENCE 26 AA: 2873 MW: 3776 CN:

 WAREHOUSE

 (TM)

 Release 3.1A John F. Collins, Biocomputing Research Unit,
 Copyright: (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd
 MFASTch_pp protein - protein database search, using Smith-Waterman algorithm
 Run On: Mon Jun 19 16:15:28 2000: Maspar time 5.59 Seconds
 Tabular output not generated.
 168.815 Million cell updates/sec

Title: >US-09-142-524A-12
 Description: (1-20) from US09.142524A.pep
 Perfect Score: 125
 Sequence: 1 SSGKNEGNYNNNEAFKVE 20

Scoring table: PAM 150
 Gap 15

Searched: 14280 seqs, 47172406 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database:

pir62
 1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 26.871; Variance 38.428; scale 0.699

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

NORMALIES

| Result No. | Query | Match | Length | ID | Description | Pred. No. |
|------------|-------|-------|--------|----|--|-----------|
| 1 | 93 | 74.4 | 374 | 2 | JC2124 major allergen Cry j I precursor (clone pCCI-15) - Japanese cedar | 5.75e-07 |
| 2 | 93 | 74.4 | 374 | 2 | JC2123 major allergen Cry j I precursor (clone pCCI-15) - Japanese cedar | 5.75e-07 |
| 3 | 69 | 54.4 | 623 | 2 | A45050 transketolase (EC 2.2.1.2) - human | 9.85e-02 |
| 4 | 63 | 50.4 | 397 | 2 | A66700 probable membrane protein - human | 8.72e-01 |
| 5 | 60 | 48.0 | 803 | 1 | RG8YH6 regulatory protein SW 3.08e-00 | 3.08e-00 |
| 6 | 59 | 47.2 | 264 | 2 | C71259 probable flagellar basal body | 4.64e-00 |
| 7 | 59 | 47.2 | 298 | 2 | T15906 hypothetical protein | 4.64e-00 |
| 8 | 59 | 47.2 | 441 | 2 | A40550 peroxisomal membrane protein | 4.64e-00 |
| 9 | 58 | 46.4 | 80 | 2 | G64030 hypothetical protein | 6.97e-00 |
| 10 | 58 | 46.4 | 643 | 2 | S36017 finger protein transmembrane | 6.97e-00 |
| 11 | 57 | 45.6 | 186 | 2 | T03992 hypothetical protein | 1.04e-01 |
| 12 | 57 | 45.6 | 265 | 2 | E70196 flagellar basal body | 1.04e-01 |
| 13 | 57 | 45.6 | 344 | 2 | A47025 abortive phage resist | 1.04e-01 |
| 14 | 57 | 45.6 | 344 | 2 | S27587 abic protein - Lactococcus lactis | 1.04e-01 |
| 15 | 57 | 45.6 | 470 | 2 | T05258 glycine hydroxymethyl transferase | 1.04e-01 |
| 16 | 57 | 45.6 | 526 | 2 | A34896 adenylate cyclase-ass | 1.04e-01 |
| 17 | 57 | 45.6 | 698 | 2 | F64839 yegH protein precursor | 1.04e-01 |
| 18 | 57 | 45.6 | 1315 | 1 | BTCLTN tentoxylisin (EC 3.4.1.1) | 1.04e-01 |
| 19 | 57 | 45.6 | 2206 | 2 | G71511 hypothetical protein | 1.04e-01 |
| 20 | 56 | 44.8 | 163 | 1 | B39415 fibrial protein smfE | 1.55e-01 |
| 21 | 56 | 44.8 | 321 | 2 | T13539 hypothetical protein | 1.55e-01 |
| 22 | 56 | 44.8 | 1365 | 1 | BVBYSK killer toxin resist | 1.55e-01 |
| 23 | 55 | 44.0 | 260 | 1 | XMBEFG flagellar basal body | 2.29e-01 |

| | | | | | | |
|----|----|------|------|---|--|----------|
| 24 | 55 | 44.0 | 295 | 2 | A42851 histamine N-methyltransferase (EC 1.1.1.18) | 2.29e-01 |
| 25 | 55 | 44.0 | 482 | 2 | S15748 hypothetical protein | 2.29e-01 |
| 26 | 55 | 44.0 | 591 | 2 | A61496 hypothetical protein | 2.29e-01 |
| 27 | 55 | 44.0 | 514 | 2 | C71156 hypothetical protein | 2.29e-01 |
| 28 | 55 | 44.0 | 568 | 2 | C24735 hypothetical protein | 2.29e-01 |
| 29 | 55 | 44.0 | 815 | 2 | T08400 hypothetical protein | 2.29e-01 |
| 30 | 55 | 44.0 | 888 | 2 | A38199 p101 protein precursor | 2.29e-01 |
| 31 | 55 | 44.0 | 1039 | 2 | E73741 hypothetical protein | 2.29e-01 |
| 32 | 54 | 43.2 | 161 | 2 | F72631 hypothetical protein | 3.38e-01 |
| 33 | 54 | 43.2 | 215 | 2 | B12905 plakoglobin, desmosomal | 3.38e-01 |
| 34 | 54 | 43.2 | 292 | 2 | JC4850 histamine N-methyltransferase | 3.38e-01 |
| 35 | 54 | 43.2 | 298 | 2 | F69875 pyrimidine-thiazine | 3.38e-01 |
| 36 | 54 | 43.2 | 331 | 2 | B55552 hec-1-hydrobenzylidene | 3.38e-01 |
| 37 | 54 | 43.2 | 337 | 2 | D72084 hec-1-hydrobenzylidene | 3.38e-01 |
| 38 | 54 | 43.2 | 481 | 2 | S57294 serine-type carboxypeptidase | 3.38e-01 |
| 39 | 54 | 43.2 | 621 | 2 | S35932 plakoglobin - mouse | 3.38e-01 |
| 40 | 54 | 43.2 | 702 | 2 | S62164 DMS1 protein - yeast | 3.38e-01 |
| 41 | 54 | 43.2 | 737 | 2 | S63453 probable PNA helicase | 3.38e-01 |
| 42 | 54 | 43.2 | 744 | 2 | A32915 plakoglobin, desmosomal | 3.38e-01 |
| 43 | 54 | 43.2 | 774 | 2 | S55133 hypothetical protein | 3.38e-01 |
| 44 | 54 | 43.2 | 1308 | 2 | E71642 probable membrane associated | 3.38e-01 |
| 45 | 54 | 43.2 | 2819 | 2 | T09610 probable chloroquine | 3.38e-01 |

AL INMENS

RESULT 1
 ENTRY JC2124 *type complete
 TITLE major allergen Cry j I precursor (clone pCCI-15) - Japanese cedar
 ORGANISM *formal name Cryptomeria japonica *common name Japanese cedar
 DATE 14-Jul-1994 *sequence revision 14-Jul-1994 *text change 26-Aug-1999
 ACCESSION JC2124
 REFERENCE JC2123
 *authors Sone, T.; Komiya, N.; Shimizu, K.; Kasakabe, T.; Morikubo, K.; Kino, K.
 *journal Biochem. Biophys. Res. Commun. (1994) 199:619-625
 *title Cloning and sequencing of cDNA coding for Cry j I, a major allergen of Japanese cedar pollen.
 *cross-references MIM:94183334

*accession JC2124
 *molecule_type mRNA
 **residues 1-374 **label N
 **cross-references G8126445, NID943643, P124006973, P124006974
 **experimental_source pollen
 **note The authors described carbohydrate binding site for histamine 474

CLASSIFICATION *supramolecular protein lysine 1454
 glycoprotein; pollen
 KEYWORDS
 FEATURE
 1-21
 22-374
 *domain signal sequence *status predicted *label SIGN
 *product major allergen Cry j I (clone pCCI-15) *status predicted *label MATN
 158,191,293,354 *binding_site carbohydrate (Asn) (covalent) *status predicted
 SUMMARY
 *length 374 *molecular-weight 40702 *checksum 3692
 Query Match 74.4% Score 93; DB 2; Length 374;
 Best Local Similarity 70.0% Pred. No. 5.75e-07;
 Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 332 SSGKNEGNYNNNEAFKVE 351
 ||||| ||||| ||||| |||||
 QY 1 SSGKNEGNYNNNEAFKVE 20

RESULT 2
 ENTRY JC2123 *type complete
 TITLE major allergen Cry j I precursor (clone pCCI-2-2) - Japanese cedar
 ORGANISM *formal name Cryptomeria japonica *common name Japanese cedar
 DATE 14-Jul-1994 *sequence revision 14-Jul-1994 *text change


```

ALTERNATE_NAMES  protein U9470.8; protein YLR182W
ORGANISM          Saccharomyces cerevisiae
DATE              31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change
                  18-Jun-1999
ACCESSIONS       S33161; S51427
REFERENCE        Breeders, J.; Nasmyth, K.
                  Nature (1987) 329:651-654
                  Similarity between cell-cycle genes of budding yeast and
                  fission yeast and the Notch gene of Drosophila.
#cross-references MIM:88014241
#accession        S33161
#molecule_type   DNA
#residues         1-803 #label BRE
#cross-references EMBL:X06238; NID:g4597; PIDN:CAA29581.1; PID:g4598
REFERENCE        S51414
                  Wohldmann, P.
                  submitted to the EMBL Data Library, November 1994
#description      The sequence of S. cerevisiae cosmid 9470.
#accession        S51427
#molecule_type   DNA
#residues         1-803 #label WOH
#cross-references EMBL:U17246; NID:g571192; PIDN:AAB67450.1;
                  PID:g577200; MIPS:YLR182W
GENETICS
#gene            SGD:SWI6; SDS11; PSL8
#map_position     12R
#cross-references SGD:S0004172; MIPS:YLR182W
CLASSIFICATION   #superfamily regulatory protein SWI6; ankyrin repeat homology
KEYWORDS         DNA binding; transcription regulation
FEATURE          317-349
                  #domain ankyrin repeat homology #label AN1\
                  #domain ankyrin repeat homology #label AK2
                  #length 803 #molecular-weight 90559 #checksum 7241
SUMMARY
Query Match      48.08; Score 60; DB 1; Length 803;
Best Local Similarity 45.08; Pred. No. 3,38e+00;
Matches          9; Conservative 6; Mismatches 5; Indels 0; Gaps 0;
DB              651 SSGIDWSSVDADPEKVE 670
CY              1 SSGKNEGNIYNNNEAFKVE 20

RESULT          6
ENTRY           671259 #type complete
TITLE           Probable flagellar basal-body rod protein (flgB1) - syphilis
                Spirochete
ORGANISM        Treponema pallidum subsp. pallidum #common_name
                syphilis spirochete
DATE            24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change
ACCESSIONS      C71259
REFERENCE        Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, G.;
                  Sutton, G.G.; Dodson, R.; Winn, M.; Hickey, E.K.; Clayton,
                  R.; Ketchum, K.A.; Sodergren, E.; Hardham, J.M.; McLeod,
                  M.P.; Salzberg, S.; Peterson, J.; Khalak, H.; Richardson,
                  D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDonald,
                  L.; Artlich, P.; Bowman, C.; Cotton, M.D.; Fujii, C.;
                  Garland, S.; Hatch, B.; Horst, K.; Roberts, K.; Watthey,
                  L.; Weidman, J.; Smith, H.O.; Venter, J.C.
                  Science (1998) 281:375-388
#journal         Complete genome sequence of Treponema pallidum, the syphilis
                  spirochete.
#accession       C71259
#status          preliminary; nucleic acid sequence not shown;
                  translation not shown;
#molecule_type DNA
#residues        1-264 #label COL
#cross-references GB:AE001264; GB:AE000520; NID:g3323278;
                  PIDN:AAE659.5.1; PID:g3323281

```

```

#experimental_source strain Nichols
GENETICS
#gene            TP0960
CLASSIFICATION   #superfamily rod protein flgF
SUMMARY          #length 264 #molecular-weight 28597 #checksum 5918
Query Match      47.28; Score 59; DB 2; Length 264;
Best Local Similarity 40.08; Pred. No. 4,54e+00;
Matches          6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
DB              109 DGTAYATGSGSEKVD 123
CY              6 EGTNTYNNNEAFKVE 20

RESULT          7
ENTRY           T15906 #type complete
TITLE           Hypothetical protein E0416.3 - Caenorhabditis elegans
ORGANISM        Caenorhabditis elegans
DATE            20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change
ACCESSIONS      T15906
REFERENCE        218427
                  Pauley, A.
                  submitted to the EMBL Data Library, June 1995
#description      The sequence of C. elegans cosmid E0416.
#accession        T15906
#status          preliminary; translated from GB/EMBL/DDBJ
#molecule_type   DNA
#residues         1-238 #label PAU
#cross-references EMBL:U28943; NID:g861333; PID:g661336;
                  PIDN:AAA68158.1; CDSF:E04F6.3
#experimental_source strain Bristol N2
GENETICS
#gene            GESP:E04F6.3
#introns         30/1; 77/2; 236/3
SUMMARY          #length 298 #molecular-weight 32721 #checksum 4744
Query Match      47.28; Score 59; DB 2; Length 298;
Best Local Similarity 36.88; Pred. No. 4,64e+00;
Matches          7; Conservative 7; Mismatches 5; Indels 0; Gaps 0;
DB              32 ARAKEDLCYVENHEDFKV 50
CY              1 SSGKNEGNIYNNNEAFKV 19

RESULT          8
ENTRY           A40550 #type complete
TITLE           Peroxisomal membrane protein, PAS3 - Yeast (Saccharomyces
                cerevisiae)
ALTERNATE_NAMES PAS3 protein; protein YDR129c
ORGANISM        Saccharomyces cerevisiae
DATE            17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change
ACCESSIONS      A40550; S59794
REFERENCE        Hoehfeld, J.; Veenhuis, M.; Kunau, W.H.
                  J. Cell Biol. (1991) 114:1167-1178
#journal         PAS3, a Saccharomyces cerevisiae gene encoding a peroxisomal
                  integral membrane protein essential for peroxisome
                  biogenesis.
#cross-references MIM:91373453
#accession       A40550
#molecule_type   DNA
#residues        1-441 #label HOE
#cross-references GB:X58407; NID:g4102; PID:g4103
REFERENCE        559783
                  Du, Z.
                  submitted to the EMBL Data Library, July 1995
#description      The sequence of S. cerevisiae cosmid 9798.
#accession       S59794
#molecule_type   DNA

```

```

**residues      1-441 #label D2J
**cross-references EMBL:U032517; NID:q914989; PID:q15004; MIPS:YR329C
**experimental_source strain S288C (AB972)

GENETICS
#gene          SGD:PEX3; PAS3
#map_position 4P
#cross-references MIPS:YR329C; SGD:S06C2737
#feature       peroxisome; transmembrane protein
#keywords      23-39
#domain transmembrane #status predicted #label IMN
#domain transmembrane #status predicted #label IM2
#length 441 #molecular_weight 50675 #checksum 156

Query Match      47.2% Score 59; DB 2; Length 441;
Best Local Similarity 46.7%; Pred. No. 4,64e-00;
Matches          7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

DB 212 SKKEGEVINEQAF 226
      |||||
QY 3 KNEGTINYNNEAF 17

#result 7
#entry 364030 #type complete
#title Hypothetical protein H1458 - Haemophilus influenzae (strain
#organism H. K120;
#formal_name Haemophilus influenzae
#date 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change
19-Jun-1998
#accessions Q64330
#reference A04000
#authors Kirschmann, R.D.; Adams, M.D.; White, O.; Clayton, F.A.;
Kirkness, E.F.; Korfavade, A.R.; Bult, C.; Tomb, J.F.;
Dougherty, B.A.; Merrick, J.M.; McKenney, K.; Sutton, G.;
FitzHugh, W.; Fields, C.; Gocayne, J.D.; Scott, J.;
Shiley, R.; Liu, L.; Glodek, A.; Kelley, J.M.; Weidman,
J.F.; Phillips, C.A.; Spriggs, T.; Hedblom, E.; Cotton,
M.D.; Gierbach, T.R.; Hanna, M.G.; Nguyen, D.T.; Saudek,
D.M.; Brandon, R.C.; Fife, L.D.; Fritchman, J.L.; Furmann,
J.L.; Geohagen, N.S.M.; Gnehm, C.L.; McDonald, L.A.;
Small, K.V.; Fraser, C.M.; Smith, R.O.; Venter, J.C.
#journal Science (1995) 269:496-512
#title Whole-genome random sequencing and assembly of Haemophilus
influenzae Rd.
#cross-references M010:Y5450630
#accession G54030
#status nucleic acid sequence not shown; translation not shown
#molecule_type DNA
#residues 1-450 #label 100K
#cross-references G54030; G54030; G54030; PID:q1574300;
H1458

GENETICS
#start_codon GTG
#summary #length 80 #molecular_weight 9400 #checksum 6360

Query Match      46.4% Score 58; DB 2; Length 80;
Best Local Similarity 42.1%; Pred. No. 6,97e-00;
Matches          8; Conservative 7; Mismatches 4; Indels 0; Gaps 0;

DB 28 NKNSTNYGNKSLIE 46
      |||||
QY 2 SKNEGTINYNNEAF 20

#result 10
#entry S36017 #type complete
#title finger protein tramtrack-p69 (ttk) - fruit fly (Drosophila
melanogaster)
#organism Drosophila melanogaster
#date 18-Jan-1995 #sequence_revision 13-Jan-1995 #text_change
24-Sep-1998
#accessions S36017; S10881
#reference S36017

```

```

#authors Montell, C.
#submission Submitted to the EMBL Data Library, April 1993
#accession S36017
#status preliminary
#molecule_type mRNA
#residues 1-643 #label M0N
#cross-references EMBL:X71627; NID:q297496; PID:q29749
#reference S10881
#authors Harrison, S.D.; Travers, A.A.
#journal EMBO J. (1990) 9:207-216
#title The tramtrack gene encodes a Drosophila finger protein that
interacts with the 1b2 transcriptional regulatory region
and shows a novel embryonic expression pattern.
#cross-references M010:50107945
#accession S10881
#status preliminary
#molecule_type DNA
#residues 3-643 #label HAP
#cross-references EMBL:X71627; NID:q29749; PID:q29749
#note the authors translated the codon GGC for residue 73 as
phe

GENETICS
#gene FlyBase:ttk
#cross-references FlyBase:FBgn0000870
#classification #superfamily PO2 domain family
#keywords DNA binding; nucleos; transcription regulation
#feature 19-120
#summary #length 643 #molecular_weight 68770 #checksum 6107

Query Match      46.4% Score 58; DB 2; Length 643;
Best Local Similarity 54.5%; Pred. No. 6,97e-00;
Matches          6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

DB 446 DYNSSND3FKL 456
      |||||
QY 9 NYNNNEAFKV 19

#result 11
#entry T03992 #type fragment
#title Hypothetical protein T03992 - Arabidopsis thaliana
(fragment)
#organism Arabidopsis thaliana #common_name mouse ear
cress
#date 30-Apr-1994 #sequence_revision 10-Apr-1994 #text_change
10-Apr-1994
#accessions T03992
#reference T03992
#authors Bewick, M.; Murphy, G.; Atthey, J.; Hudson, S.; Hamblin, J.;
Mowes, H.W.; Mayer, K.F.X.; Schaeffer, G.
#submission Submitted to the Protein Sequence Database, March 1993
#accession T03992
#molecule_type DNA
#residues 1-185 #label BEV
#cross-references EMBL:AL042481
#experimental_source cultivar Col-0;mb: BAC clone TSL19

GENETICS
#map_position 4
#note Intronic positions not resolved; TSL19.10
#summary #length 185 #checksum 902

Query Match      45.6% Score 57; DB 2; Length 185;
Best Local Similarity 40.0%; Pred. No. 1,04e-01;
Matches          5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

DB 102 KNEGVGYNDKIFR 116
      |||||
QY 4 KNEGTINYNNEAF 19

#result 12
#entry E70196 #type complete

```

 W P S R E A

 (TM)

Release 3.1A John F. Collins, Biocomputing Research Unit,
 Copyright (C) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Wpsrch_pp protein - protein database search, using Smith-Waterman algorithm
 Run on: Mon Jun 19 16:14:36 2000; MasPar time 3.71 Seconds
 Tabular output not generated. 164.264 Million cell updates/sec

Title: >US-09-142-524A-12
 Description: (1-20) from US09142524A.pep
 Perfect Score: 125
 Sequence: 1 SSGKNEGNIYNNNEAFKVE 20

Scoring table: PAM 150
 Gap 15

Searched: 83957 seqs, 30454573 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: swiss-prot38
 1:swissprot

Statistics: Mean 27.434; Variance 35.196; scale 0.780

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed.
 and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description | Pred. No. |
|------------|-------|---------------|--------|-------|-------------|-----------|
| 1 | 93 | 74.4 | 374 | 1 | SBP_CRYJA | 4.55e+08 |
| 2 | 68 | 54.4 | 623 | 1 | TKT_HUMAN | 2.31e-02 |
| 3 | 64 | 51.2 | 623 | 1 | TKT_MOUSE | 1.57e-01 |
| 4 | 64 | 51.2 | 623 | 1 | TKT_RAT | 1.57e-01 |
| 5 | 60 | 48.0 | 803 | 1 | SW16_YEAST | 9.94e-01 |
| 6 | 60 | 48.0 | 843 | 1 | CYAA_YEAST | 9.94e-01 |
| 7 | 59 | 47.2 | 441 | 1 | PEX3_YEAST | 1.56e+00 |
| 8 | 58 | 46.4 | 80 | 1 | YK99_HAEN | 2.43e+00 |
| 9 | 58 | 46.4 | 641 | 1 | TTKB_DROME | 2.43e+00 |
| 10 | 57 | 45.6 | 265 | 1 | FLGG_BORBU | 3.76e+00 |
| 11 | 57 | 45.6 | 344 | 1 | ABIC_LACIA | 3.76e+00 |
| 12 | 57 | 45.6 | 526 | 1 | CAP_YEAST | 3.76e+00 |
| 13 | 57 | 45.6 | 698 | 1 | YMC4_ECOLI | 3.76e+00 |
| 14 | 57 | 45.6 | 713 | 1 | PALY_DICHA | 3.76e+00 |
| 15 | 57 | 45.6 | 1116 | 1 | MHK1_SCHPO | 3.76e+00 |
| 16 | 57 | 45.6 | 1314 | 1 | TEXI_CLOTE | 3.76e+00 |
| 17 | 56 | 44.8 | 1365 | 1 | KRS5_YEAST | 5.79e+00 |
| 18 | 56 | 44.8 | 3555 | 1 | YAKB_SCHPO | 5.79e+00 |
| 19 | 55 | 44.0 | 260 | 1 | FLGG_SALTY | 8.87e+00 |
| 20 | 55 | 44.0 | 294 | 1 | HMT_RAT | 8.87e+00 |
| 21 | 55 | 44.0 | 378 | 1 | KLP2_BOMMO | 8.87e+00 |
| 22 | 55 | 44.0 | 482 | 1 | NIFD_AZOBR | 8.87e+00 |
| 23 | 55 | 44.0 | 491 | 1 | YF86_METJA | 8.87e+00 |

| | | | | | | | |
|----|----|------|------|---|------------|-------------------------|----------|
| 24 | 54 | 43.2 | 159 | 1 | GLBR_CAEEL | POTATIVE GLOBIN-LIKE P | 1.35e+01 |
| 25 | 54 | 43.2 | 292 | 1 | HMT_HUMAN | HISTAMINE N-METHYLTRAN | 1.35e+01 |
| 26 | 54 | 43.2 | 298 | 1 | APBA_BACSU | PROBABLE 2-DEHYDROPIANT | 1.35e+01 |
| 27 | 54 | 43.2 | 315 | 1 | OMPP_ECOLI | OUTER MEMBRANE PROTEAS | 1.35e+01 |
| 28 | 54 | 43.2 | 331 | 1 | NAV1_PSEPU | TRANS-O-HYDROXYBENZYLIT | 1.35e+01 |
| 29 | 54 | 43.2 | 334 | 1 | NAV1_PSEPU | HOLLIDAY JUNCTION DNA | 1.35e+01 |
| 30 | 54 | 43.2 | 334 | 1 | NAV1_PSEPU | HOLLIDAY JUNCTION DNA | 1.35e+01 |
| 31 | 54 | 43.2 | 337 | 1 | YU45_MYCCE | HYPOTHETICAL LIPOPROTE | 1.35e+01 |
| 32 | 54 | 43.2 | 483 | 1 | CATA_HELAN | CATALASE (EC 1.11.1.6) | 1.35e+01 |
| 33 | 54 | 43.2 | 492 | 1 | PLAK_MOUSE | JUNCTION PLAKOGLOBIN | 1.35e+01 |
| 34 | 54 | 43.2 | 621 | 1 | EXOL_YEAST | EXONUCLEASE 1 (EXO 1) | 1.35e+01 |
| 35 | 54 | 43.2 | 732 | 1 | SUV3_YEAS | MITOCHONDRIAL ATP-DEPE | 1.35e+01 |
| 36 | 54 | 43.2 | 737 | 1 | PLAK_HUMAN | JUNCTION PLAKOGLOBIN | 1.35e+01 |
| 37 | 54 | 43.2 | 743 | 1 | YMA4_YEAS | HYPOTHETICAL 87.1 KD F | 1.35e+01 |
| 38 | 54 | 43.2 | 774 | 1 | YMA4_HUMAN | HYPOTHETICAL PROTEIN K | 1.35e+01 |
| 39 | 54 | 43.2 | 1942 | 1 | RNP_EGJIG | RIBONUCLEASE PANCHEAT | 2.04e+01 |
| 40 | 53 | 42.4 | 119 | 1 | HS18_CLOAB | 19 KD HEAT SHOCK PROTE | 2.04e+01 |
| 41 | 53 | 42.4 | 151 | 1 | YAKB_SCHPO | HYPOTHETICAL 27.9 KD P | 2.04e+01 |
| 42 | 53 | 42.4 | 251 | 1 | YU45_MYCCE | HYPOTHETICAL PROTEIN G | 2.04e+01 |
| 43 | 53 | 42.4 | 540 | 1 | ENV_HVIA2 | ENVELOPE POLYPROTEIN G | 2.04e+01 |
| 44 | 53 | 42.4 | 855 | 1 | YOG7_CAEEL | HYPOTHETICAL 375.7 KD | 2.04e+01 |
| 45 | 53 | 42.4 | 3343 | 1 | YOG7_CAEEL | HYPOTHETICAL 375.7 KD | 2.04e+01 |

ALIGNMENTS

| RESULT | ID | SBP_CRYJA | STANDARD | PROJ | 374 AA |
|--------|--|-----------------------------------|----------|------|--------|
| AC | P18632 | | | | |
| DT | 01-NOV-1990 | (Rel. 16, Created) | | | |
| DT | 01-NOV-1995 | (Rel. 32, Last sequence update) | | | |
| DT | 15-DEC-1998 | (Rel. 37, Last annotation update) | | | |
| DE | SUGI BASIC PROTEIN PRECURSOR (SBP) (MAJOR ALLERGEN CRY J 1) (CRY J 1) | | | | |
| OS | Cryptomeria japonica (Japanese cedar) | | | | |
| OC | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; euphyllophytes; Spermatophyta; Coniferopsida; Coniferales; Taxodiaceae; Cryptomeria | | | | |
| OC | [1] | | | | |
| RN | SEQUENCE FROM N.A., AND PARTIAL SEQUENCE. | | | | |
| RP | TSSTVE-POLLEN: | | | | |
| RX | MEDLINE: 94181234. | | | | |
| RA | Sone I., Komiyama N., Shimizu K., Sakai T., Morikubo K. | | | | |
| RA | Kino K. | | | | |
| RI | Isolating and sequencing of cDNA clones for the major allergen of Japanese cedar pollen. | | | | |
| RI | Biochem. Biophys. Res. Commun. 199, 119-124 (1994) | | | | |
| RN | [2] | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RC | TISSUE=POLLEN: | | | | |
| RA | Namba M., Kurose M., Torigoe K., Fukuda S., Kurimoto M. | | | | |
| RA | Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases. | | | | |
| RL | [3] | | | | |
| RP | SEQUENCE OF 22-41. | | | | |
| RC | TISSUE=POLLEN: | | | | |
| RX | MEDLINE: 89031257. | | | | |
| RA | Taniai M., Ando S., Usui M., Kurimoto M., Sakaguchi M., Inouye S., Matsui T. | | | | |
| RA | "N-terminal amino acid sequence of a major allergen of Japanese cedar pollen (Cry j 1)". | | | | |
| RL | FEBS Lett. 239:329-332(1988). | | | | |
| RP | CARBOHYDRATES. | | | | |
| RA | MEDLINE: 95003748. | | | | |
| RA | Hijikata A., Matsumoto I., Kojima K., Ogawa H. | | | | |
| RT | "Antigenicity of the oligosaccharide moiety of the Japanese cedar (Cryptomeria japonica) pollen allergen, Cry j1". | | | | |
| RL | Int. Arch. Allergy Immunol. 105:198-202(1994). | | | | |
| CC | !- PTM: CONTAINS FUCOSEXYLOSE-CONTAINING N-LINKED OLIGOSACCHARIDES. | | | | |
| CC | !- DISEASE: THIS PROTEIN IS THE MAJOR ALLERGEN OF JAPANESE CEDAR POLLEN, THE MOST COMMON POLLEN ALLERGEN IN JAPAN. | | | | |
| CC | !- MISCELLANEOUS: THE SEQUENCE OF CRY J 1 FORM A IS SHOWN HERE. FORM B DIFFERS IN SIX POSITIONS. | | | | |
| CC | !- SIMILARITY: BELONGS TO THE OLYSACCHARIDE LYASE FAMILY 1. | | | | |

AMB A 1/AMB A 11/CRY J 1 SUBFAMILY

 This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL: D26544; AAA05542.1; 1;
 EMBL: D26545; AAA05543.1; 1;
 EMBL: D34639; AAA07020.1; 1;
 PIR: A44773; A44773;
 PRAM: PF00544; pectylase; 1;
 PRINTS: PR08007; AMBALERGEN;
 KW Allergen; Glycoprotein; Multigene family; Signal;
 FT SIGNAL 1 22
 FT CHAIN 22 374
 FT VARIAT 14 143
 FT VARIAT 202 202
 FT VARIAT 221 221
 FT VARIAT 358 358
 FT VARIAT 361 361
 FT CARBOHYD 154 158
 FT CARBOHYD 191 191
 FT CARBOHYD 293 293
 SEQUENCE 374 AA; 40645 MW; 74AB25950248F56F CRC64;
 Query Match: 74.48; Score 93; DR 1; Length 374;
 Best Local Similarity 70.0%; Pred No. 4.05e-08;
 Matches 14; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

 1 SSGKVEGNIYIKKEAFNVE 351
 1 SSGKNEGNIYNNKEAFNVE 20

 RESULT 2
 ID TKI HUMAN STANDARD; PRI: 623 AA;
 AC P29410;
 DT 01-DEC-1992 (Ref. 24, Created);
 DI 01-OCT-1996 (Ref. 34, Last sequence update);
 DE TRANSKETOLASE (EC 2.2.1.1) (TK);
 GN TKI (CR TKI);
 OS Homo sapiens (Human);
 OC Eukaryota; Metazoa; Chordata; Gibbata; Vertebrata; Mammalia;
 OC Eutheria; Primates; Catarrhini; Hominoidea; Homo;
 EC [1];
 SEQUENCE FROM N.A.
 TISSUE=LIVER;
 RX MEDLINE; 93123263;
 RA McCool B.A., Pionk S.G., Martin P.R., Singleton C.K.;
 RT "Cloning of human transketolase cDNAs and comparison of the
 RT nucleotide sequence of the coding region in Wernicke-Korsakoff and
 RT non-Wernicke-Korsakoff individuals";
 PL J. Biol. Chem. 268:1397-1404(1993);
 RN [2];
 RP SEQUENCE FROM N.A.
 RA Carr F.H., Shew K.F.R.E., Szabo P., Plase J.P.;
 RI Submitted (SEP-1993) to the EMBL/GenBank/DBJ databases;
 RN [3];
 RP SEQUENCE FROM N.A.
 RA Scherk G., Layfield R., Candy J.M., Duggleby R.G., Nixon P.F.;
 RI Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases;
 RN [4];
 RP SEQUENCE OF 224-623 FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE; 92231878.

RA Abedinia M., Layfield R., Jones S.M., Nixon P.F., Mattick J.S.;
 RT "Nucleotide and predicted amino acid sequence of a cDNA clone
 RT encoding part of human transketolase";
 RL Biochem. Biophys. Res. Commun. 183:1159-1166(1992);
 CC 3-; CATALYTIC ACTIVITY: SEDOHEPTULOSE 7-PHOSPHATE + D-GLYCERALDEHYDE
 CC 3-PHOSPHATE + D-RIBOSE 5-PHOSPHATE -> D-XULOSE 5-PHOSPHATE;
 CC 3-COFACITOR: THIAMINE PYROPHOSPHATE;
 CC 3-SUBUNIT: HOMODIMER;
 CC 3-DISEASE: HAS BEEN IMPLICATED IN THE LATENT GENETIC DISEASE,
 CC WERNICKE-KORSAKOFF SYNDROME (WKS) WHICH CAUSE SPECIFIC BRAIN
 CC DAMAGE;
 CC 3-SIMILARITY: BELONGS TO THE TRANSKETOLASE FAMILY;
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 EMBL: X67588; CAA47919.1; 1;
 EMBL: L22713; AAA61222.1; 1;
 EMBL: U55017; AAA98961.1; 1;
 PIR: A45050; A45050;
 PIR: PH0845; PH0845;
 MIM: 277730;
 DR PROSITE: PS00801; TRANSKETOLASE; 1;
 DR PROSITE: PS00802; TRANSKETOLASE; 1;
 DR PRAM: PF00456; transketolase; 1;
 KW Transferrase; Thiamine pyrophosphate;
 FT CONFLICT 30 31 IT -> SS (IN REF. 1);
 FT CONFLICT 45 46 I -> V (IN REF. 1);
 FT CONFLICT 224 230 CKAFCQ -> AVQVCE (IN REF. 4);
 FT CONFLICT 426 426 F A (IN REF. 1);
 FT CONFLICT 585 587 TRL -> KTM (IN REF. 2);
 FT CONFLICT 608 623 LRLAQAQVRLTITKA -> TGPLPLKL (IN REF. 4);
 SEQUENCE 623 AA; 67877 MW; 176-89032F627125 CRC64;
 Query Match: 54.4%; Score 68; DR 1; Length 623;
 Best Local Similarity 50.0%; Pred No. 2.31e-12;
 Matches 5; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

 DB 472 TSPENALYNNNEAFV 48;
 QY 2 SSKNEGNIYNNKEAFNVE 14

 RESULT 3
 ID TKI MOUSE STANDARD; PRI: 523 AA;
 AC P40142;
 DT 01-FEB-1995 (Ref. 31, Created);
 DI 01-FEB-1995 (Ref. 31, Last sequence update);
 DI 15-FEB-2000 (Ref. 39, Last annotation update);
 DE TRANSKETOLASE (EC 2.2.1.1) (TK) (P66);
 GN TKI;
 OS Mus musculus (Mouse);
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
 OC Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus;
 EC [1];
 RP SEQUENCE FROM N.A.
 RC STRAIN=LAF1;
 RX MEDLINE; 96214928;
 RA Schimmer B.P., Tsao J., Czerwinski, W.;
 RT "Amplification of the transketolase gene in desensitization-resistant
 RT mutant B6 mouse adrenocortical tumor cells";
 RL J. Biol. Chem. 271:4993-4998(1996);
 RN [2];
 RP SEQUENCE OF 1-36 FROM N.A.
 RC STRAIN=129/SV;
 RA Salamon C., Chervenak M., Platiqorsky J., Sax C.M.;
 RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases;
 CC 3-; CATALYTIC ACTIVITY: SEDOHEPTULOSE 7-PHOSPHATE + D-GLYCERALDEHYDE

DB 472 TSREPENAIYSNNEQFV 489
QY 2 SKNEGTYNNNEAFKV 19

RESULT 4
ID TKT RAI STANDARD: PRT: 623 AA.
IP P50137.
DT 01-OCT-1996 (Rel. 34, Created)
DI 01-OCT-1996 (Rel. 34, Last sequence update)
DE TRANSKETOLASE (EC 2.2.1.1) (TK).
GN TK.
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia;
OC Eutheria; Rodentia; Sciurognathia; Muridae; Murinae; Rattus.
RN [1]
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LIVER;
RA Kim S., Kim B., Jeng J., Song B.,
RE Submitted (APR-1994) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: SEDUCHEPILULOSE 7-PHOSPHATE - D-XYLULOSE 5-PHOSPHATE
CC -!- 3-PHOSPHATE - D-RIBOSE 5-PHOSPHATE + D-XYLULOSE 5-PHOSPHATE
CC -!- COFACTOR: THIAMINE PYROPHOSPHATE.
CC -!- SUB-UNIT: HOMODIMER (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE TRANSKETOLASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U05809; AAC52443.1; .
CC EMBL: U05889; AAC53570.1; .
CC SWISS-2DPAGE: P40142; M0USE.
CC MGD: MG:105992; TKT.
CC PROSITE: PS00801; TRANSKETOLASE.1; 1.
CC PROSITE: PS00802; TRANSKETOLASE.2; 1.
CC PFAM: PF00456; Transketolase; 1.
CC Transferase: Thiamine pyrophosphate.
KW TRANSFERASE; THIAMINE PYROPHOSPHATE.
SQ SEQUENCE 623 AA: 67633 MW: 870045ADS558FA09 CRC64:
Query Match 51.2% Score 64: DB 1: Length 623;
Best Local Similarity 44.4% Pred. NC: 1.57e-01;
Matches 8: Conservative 5: Mismatches 5: Indels 0: Gaps 0:

DB 472 TSREPENAIYSNNEQFV 489
QY 2 SKNEGTYNNNEAFKV 19

RESULT 5
ID SW16_YEARST STANDARD: PRI: 803 AA
AC P09959;
DT 01-MAR-1989 (Rel. 10, Created)
DI 01-MAR-1989 (Rel. 10, Last sequence update)
DE REGULATORY PROTEIN SW16 (CELL-CYCLE BOX FACTOR, CHAIN SW16) (TRANS-
GN ACTING ACTIVATOR OF HO ENDOGENOUS GENE) (MFE SUBUNIT P30).
DE SW16 OR YLR182W OR L947C8.
OS Saccharomyces cerevisiae (Baker's Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales.
CC Saccharomycetaceae; Saccharomyces.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 880:4241.
RA Breeden L., Nasmyth K.,
RT "Similarity between cell-cycle genes of budding yeast and tissue
RC yeast and the Notch gene of Drosophila."
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
RA Favella A., Fulton L., Gattung S., Greco T., Kirsten T.,
RA Kucaba T., Hallsworth K., Hawkins J., Hillier S., Jier M.,
RA Johnson D., Johnston L., Landston Y., Sattelle P., Le J.,
RA Maris E., Menezes S., Miller N., Nhan M., Pauley A., Peluso D.,
RA Rifkin L., Riles L., Tatch A., Trevasakis E., Vignati D.,
RA Wilcox L., Wolfdran P., Vaudin M., Wilson R., Waterston K.,
RL Submitted (NOV-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP CHARACTERIZATION.
RX MEDLINE: 92301535.
RA Dirick L., Moll T., Auer H., Nasmyth K.,
RT "A central role for Sw16 in modulating cell cycle Start-specific
RN transcription in yeast."
RL Nature 357:508-513(1992).
CC -!- FUNCTION: PART OF A COMPLEX INVOLVED IN CELL-CYCLE-DEPENDENT
CC TRANSCRIPTION. SW16 AND SW14 ARE REQUIRED FOR FORMATION OF THE
CC CELL-CYCLE BOX FACTOR-DNA COMPLEX. THE REPEATED ELEMENT IN THE
CC UPSTREAM REGION OF HES1 (5'-GAGGAGAAA-3') IS CALLED THE CELL-CYCLE
CC BOX (CCB).
CC -!- SUB-UNIT: MBF CONTAINS SW16 AND MBP1. MBF CONTAINS SW16 AND SW14.
CC -!- SIMILARITY: STRONG, TO S. PUMBE CUBIL.
CC -!- SIMILARITY: CONTAINS 2 ANK REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X05338; CAA29581.1; .
CC EMBL: U1246; AAB67460.1; .
CC PIR: S03161; RBYW6.
CC TRANSFAC: T00096; .
CC TRANSFAC: T01013; .
CC SGD: L0002254; SW16.
CC PFAM: PF00023; ank; 2.
CC Transcription regulation: DNA-binding; ANK repeat; Repeat.
FT DOMAIN 311 495 2 X ANK MOTIF REPEATS.
FT REPEAT 311 343 ANK MOTIF 1.
FT REPEAT 453 495 ANK MOTIF 2.
FT DOMAIN 631 640 GLU-RICH (ACIDIC).
SQ SEQUENCE 803 AA: 90559 MW: 9B317FCAGEC493C CRC64;

Query Match 48.0% Score 60; DB 1; Length 803;
 Best Local Similarity 45.0%; Pred. No. 9,94e-01;
 Matches 9; Conservative 6; Mismatches 5; Indels 0;

2b 651 SSGIDMSSEYDADEPFKE 670

QY 1 SSGKNEGNIYNNNEAFKVE 20

RESULT 6
 ID CYAA_AERHY STANDARD; PRI: 843 AA.

AC Q59119;

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DI 01-NOV-1997 (Rel. 35, Last sequence update)

DE ADENYLATE CYCLASE (EC 4.6.1.1) (AIP PYROPHOSPHATE-LYASE); (ADENYLYL

DE CYCLASE);

GN CYA.

OS Aeromonas hydrophila

CC Bacteria; Proteobacteria; gamma subdivision: Aeromonas group;

CC Aetomonas.

CC [1]

CC SEQUENCE FROM N.A.

CC STRAIN-218; Vivars C., Glaser P., Roy A., Danchin A.

CC Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.

CC -1- CATALYTIC ACTIVITY: AIP = 3',5'-CYCLIC AMP + PYROPHOSPHATE.

CC -1- ENZYME REGULATION: THE REGULATORY DOMAIN IS INVOLVED IN THE

CC REGULATION OF CYCLASE ACTIVITY BY THE CARBON SOURCE.

CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.

CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-1 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL: X95881; CAA65130.1;

CC PROSITE: PSC1042; ADENYLATE CYCLASE_1; 1;

CC PROSITE: PSC1093; ADENYLATE CYCLASE_12; 1;

CC PFAM: PF01295; Adenylate_cyclase_1;

CC Lyase; CAMP synthetase.

CC DOMAIN 543 843

CC REGULATORY.

CC SEQUENCE 843 AA; 96852 MW; 90E267AC2AB8C77 CR 64;

Query Match 48.0% Score 60; DB 1; Length 843;

Best Local Similarity 52.9%; Pred. No. 9,94e-01;

Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

2b 750 RGAGTNIYILDEANRVE 766

QY 4 KNSGGINNNNEAFKVE 20

RESULT 7
 ID PEX3_YEAST STANDARD; PRI: 441 AA.

AC P28795;

DT 01-DEC-1992 (Rel. 24, Created)

DI 01-DEC-1992 (Rel. 35, Last sequence update)

DE PEROXISOMAL MEMBRANE PROTEIN PAS3 (PEROX-N-3).

GN PEX3 OR PAS3 OR YDR329C OR D9798.15.

OS Saccharomyces cerevisiae (Baker's yeast).

CC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;

CC Saccharomycetaceae; Saccharomycetes.

CC [1]

CC SEQUENCE FROM N.A.

CC MEDLINE: 91373453.

CC Hoehfeld J., Veenhuis M., Kunau W.H.;

*PAS3, a Saccharomyces cerevisiae gene encoding a peroxisomal
 integral membrane protein essential for peroxisome biogenesis.
 J. Cell Biol. 114:1167-1178(1991).

[2]

SEQUENCE FROM N.A.

CC STRAIN-S288C / AB972;

CC Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Li Z.,

CC Faveillo A., Fulton L., Gattung S., Greco T., Kirsten J.,

CC Kucaba T., Hallsworth K., Hawkins J., Hillier L., Jier M.,

CC Johnson D., Johnston L., Langston Y., Latreille P., Le I.,

CC Mardis E., Menezes S., Miller N., Nham M., Pauley A., Peluso D.,

CC Rifkin L., Riles L., Taich, Trevasaki E., Vignati D.,

CC Wilcox L., Wolfdman P., Vaud M., Wilson R., Waterston R.,

CC submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: INVOLVED IN PEROXISOME BIOSYNTHESIS. MAY FUNCTION

CC AS A RECEPTOR PROTEIN. LACK OF THIS PROTEIN CAUSES THE

CC PEROXISOMAL-DEFICIENT PHENOTYPE AND MISLOCALIZATION IN THE

CC CYTOSOL OF PEROXISOMAL MATRIX PROTEINS.

CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. PEROXISOMAL.

CC -1- INDUCTION: BY OLEIC ACID.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL Outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL: X58407; CAA4339.1;

CC EMBL: J32517; AAB64764.1;

CC PIR: A40550; A40550.

CC SGI: L0001339; PAS3.

CC Transmembrane; Peroxisome.

CC DOMAIN 1 17 MATRIX (POTENTIAL).

CC TRANSMEM 18 39 POTENTIAL.

CC DOMAIN 40 441 CYTOPLASMIC (POTENTIAL).

CC SEQUENCE 441 AA; 56675 MW; 3CDS18356F0B3F9 CR664;

Query Match 47.2% Score 59; DB 1; Length 441;

Best Local Similarity 46.7%; Pred. No. 1,56e-00;

Matches 7; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

Db 2:2 SKKEGQEWYINQAF 226

QY 3 GKNEGNIYNNNEAF 17

RESULT 8

ID YE58_HAEIN STANDARD; PRI: 80 AA.

AC P44234;

DI 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DE HYPOTHETICAL PROTEIN HI1458.

GN HI1458.

OS Haemophilus influenzae.

CC Bacteria; Proteobacteria; gamma subdivision: Pasteurellales;

CC Haemophilus.

CC [1]

CC SEQUENCE FROM N.A.

CC STRAIN-RD / KW20;

CC MEDLINE: 95350630.

CC Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,

CC Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,

CC McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,

CC Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,

CC Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,

CC Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,

CC Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghagen N.S.M.,

CC Gnehm C., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,

CC Venter J.C.;

CC *Whole-genome random sequencing and assembly of Haemophilus

Query Match 45.6% Score 57 DB 1 Length 526;
 Best Local Similarity 40.0% Pred. No. 3,756-00;
 Matches 8: Conservative 7; Mismatches 5; Indels 0; Gaps 0;

DB 45% SIKSDSGNINVLKESANTE 478
 QY : SSGKNEGINYNNEAFKE 20

RESULT 13
 ID YMCALSCOLI STANDARD: PRT: 698 AA.
 AC P15882;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-FEB-2000 (Rel. 39, Last annotation update)
 DE HYPOTHETICAL 78.7 KD LIPOPROTEIN IN APPA-CSPH INTERGENIC REGION
 DE PRECURSOR.
 GN YMCAL.
 OS Escherichia coli.
 CC Bacteria; Proteobacteria; gamma subdivision: Enterobacteriaceae;
 CC Escherichia.
 KS [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE: 97426517.
 RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Corrado-Vides J., Glasner F.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau E., Shao Y.
 RT "The complete genome sequence of Escherichia coli K-12."
 RC Science 277:1453-1474(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE: 97061202.
 RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,
 RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
 RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
 RA Mori H., Motomura K., Nakamura Y., Nishino Y., Saito N.,
 RA Sampei G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,
 RA Yano M., Horuchi T.
 RA "A 7.8-kb DNA sequence of the Escherichia coli K-12 genome
 RT corresponding to the 12.7-28.0 min region on the linkage map."
 RC DNA RES. 3:137-155(1994).
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A LIPID
 CC ANCHOR (PROBABLE).
 CC -1- SIMILARITY: STRONG. TO EMBL YVRR.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sit.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AE000200; AAC74069.1;
 CC EMBL: D90735; BAA35749.1;
 CC EMBL: D90736; BAA36124.1;
 CC ECOGENE: E013729; YMCAL.
 CC PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 CC Hypothetical protein; Membrane; Lipoprotein; Signal.
 FT SIGNAL 1 18
 FT CHAIN 19 698
 FT LIPID 15 19
 FT ACYD DIGLYCERIDE (POTENTIAL).
 CC SEQUENCE 698 AA: 78687 MW: 37571D2C57FA4E7 CRC64;

Query Match 45.6% Score 57 DB 1 Length 698;
 Best Local Similarity 72.9% Pred. No. 3,756-00;
 Matches 8: Conservative 1; Mismatches 2; Indels 0; Gaps 0;

DB 470 IN'YNNYKFK 480

QY 8 IN'YNNNEAFK 18

RESULT 14
 ID PALLYDIGLA STANDARD: PRT: 713 AA.
 AC Q23924;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE PHENYLALANINE AMMONIA-LYASE (EC 4.3.1.5).
 OS Digitalis lanata (Foxglove).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 CC core eudicots; Asteridae; euasterids I; Lamiales; Scrophulariales;
 CC Digitalis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Thoeninger C.
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: THIS IS A KEY ENZYME OF PLANT METABOLISM CATALYZING THE
 CC FIRST REACTION IN THE BIOSYNTHESIS FROM L-PHENYLALANINE OF A WIDE
 CC VARIETY OF NATURAL PRODUCTS BASED ON THE PHENYLPROPANE SKELETON.
 CC -1- CATALYTIC ACTIVITY: L-PHENYLALANINE + TRANS-CINNAMATE -> NH(3).
 CC -1- COFACTOR: HAS A DEHYDROLANINE (DHA) IN THE ACTIVE SITE.
 CC -1- PATHWAY: KEY ENZYME OF PHENYLPROPANOID METABOLISM.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (PROBABLE).
 CC -1- SIMILARITY: BELONGS TO THE PAL / HISTIDASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sit.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AJ062221; CAA05251.1;
 CC PROSITE: PS00488; PAL_HISTIDASE; 1.
 CC PFAM: PF00221; PAL; 1.
 CC Lysase; Phenylpropanoid metabolism; Multi-gene family.
 FT ACT SITE 199 199 BY SIMILARITY
 CC SEQUENCE 713 AA: 77732 MW: 747564669D0C6A CRC64;

Query Match 45.6% Score 57 DB 1 Length 713;
 Best Local Similarity 47.0% Pred. No. 3,756-00;
 Matches 8: Conservative 4; Mismatches 5; Indels 0;

DB 60% KNASISIFCKIEAFKE 522
 QY 4 KNEGINYNNEAFKE 20

RESULT 15
 ID MKH1SCHPG STANDARD: PRT: 1116 AA.
 AC Q10407;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-FEB-2000 (Rel. 39, Last annotation update)
 DE SERINE/THREONINE-PROTEIN KINASE MKH1 (EC 2.7.1.1).
 GN MKH1 OR PMK1 OR SPAC1F3.02C.
 OS Schizosaccharomyces pombe (Fission yeast).
 CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 CC Schizosaccharomycetaceae; Schizosaccharomycetes.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SP66;
 RX MEDLINE: 97342589.
 RA Sengar A.S., Markley N.A., Marini N.J., Young D.;
 RT "Mkh1, a MEK kinase required for cell wall integrity and proper
 RT response to osmotic and temperature stress in Schizosaccharomycetes
 RT pombe.";
 RT Mol. Cell. Biol. 17:3508-3519(1997).

```

RN      (2)
RP      SEQUENCE FROM N.A.
RA      SIRAIN-972:
RA      CORCOR R., Church C.M., Barrell B.G., Rajandream M.A., Walsh S.V.:
RA      Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
RL      -!- FUNCTION: MAY REGULATE CELL MORPHOLOGY, CELL WALL INTEGRITY, SALT
CC      RESISTANCE, CELL CYCLE REENTRY FROM STATIONARY-PHASE ARREST, AND
CC      FILAMENTOUS GROWTH IN RESPONSE TO STRESS.
CC      -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC      MAP KINASE KINASE SUBFAMILY.
CC      .....
CC      THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      .....
DR      EMBL: U53872; AAB62119.1;
DR      EMBL: Z70690; CAA94523.1;
DR      HSSP: P24941; 1AQ1.
DR      PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR      PROSITE: PS00109; PROTEIN_KINASE_ST; 1.
DR      PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR      PFAM: PF00069; pkinase; 1.
KW      Transferase; Serine/threonine-protein kinase; Tyrosine-protein kinase;
KW      ATP-binding.
FT      DOMAIN 825 1094 PROTEIN KINASE.
FT      NP_BIND 831 839 ATP (BY SIMILARITY).
FT      BINDING 854 854 ATP (BY SIMILARITY).
FT      ACT_SITE 955 955 BY SIMILARITY.
SQ      SEQUENCE 1116 AA; 225132 MW; 7AFDB9EC62ED47B CRC64;

Query Match 45.6% Score 57 DB 1 Length 1116;
Best Local Similarity 40.0% Pred. No. 3.76e+00;
Matches 6; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

> 402 KADGDSIVYRKEDEK 416
> .....
> 4 KNEGNIYNNNEAFK 18

```

Search completed: Mon Jun 19 16:14:43 2000
 Job time : 7 secs.

WATERMAN

(TM)

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Jun 19 16:14:59 2000: MasPar time 9.10 Seconds
152.449 Million cell updates/sec

Incubator output not generated.

Title: >US-09-142-524A-12
Description: (1-25) from US9142524A.pep
Perfect Score: 125
Sequence: 1 SSGKNEGNIYNNNEAFKVE 20

Scoring table: PAM 150
Gap 15

Searched: 225878 seqs. 69334122 res.idues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: spltrebbl2
1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human
5:sp-invertebrate 6:sp-mammal 7:sp-mhc 8:sp-organelle
9:sp-phage 10:sp-plant 11:sp-rodent 12:sp-unclassified
13:sp-vertebrate 14:sp-virus

Statistics: Mean 26.556; Variance 35.964; scale 0.738

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description | Pred. No. |
|------------|-------|-------------|--------|-------|-----------------------------------|-----------|
| 1 | 125 | 100.0 | 375 | 10 | CHA01 PRECURSOR. | 2.25e-15 |
| 2 | 108 | 86.4 | 367 | 10 | POLLEN MAJOR ALLERGEN | 4.37e-11 |
| 3 | 69 | 55.2 | 1598 | 10 | DYNEIN HEAVY CHAIN ISO | 4.01e-02 |
| 4 | 53 | 50.4 | 397 | 3 | CHROMOSOME XV READING | 6.52e-01 |
| 5 | 51 | 48.8 | 420 | 2 | NADH OXIDASE (FRAGMENT | 1.60e-00 |
| 6 | 60 | 48.0 | 126 | 13 | TRANSKETOLASE (FRAGMENT | 2.48e+00 |
| 7 | 59 | 47.2 | 146 | 2 | 17 KD HEMAGGLUTININ CO | 3.84e+00 |
| 8 | 59 | 47.2 | 146 | 2 | Q45841 | 3.84e+00 |
| 9 | 59 | 47.2 | 264 | 2 | FLAGELLAR BASAL-BODY R | 3.84e+00 |
| 10 | 59 | 47.2 | 298 | 5 | 19C058 | 3.84e+00 |
| 11 | 58 | 46.4 | 290 | 2 | BETA-ACARASE I PRECURS | 5.92e+00 |
| 12 | 58 | 46.4 | 379 | 14 | Q9W482 | 5.92e+00 |
| 13 | 58 | 46.4 | 432 | 2 | HYPOTHETICAL 50.7 KD P | 5.92e+00 |
| 14 | 57 | 45.6 | 2206 | 5 | Q92605 | 5.92e+00 |
| 15 | 57 | 45.6 | 4550 | 5 | HYPOTHETICAL 264.1 KD | 9.07e+00 |
| 16 | 57 | 45.6 | 4981 | 5 | PFC0425W PROTEIN. | 9.07e+00 |
| 17 | 56 | 44.8 | 163 | 2 | PFC0820W PROTEIN. | 9.07e+00 |
| 18 | 56 | 44.8 | 321 | 9 | SMFE PROTEIN PRECURSOR | 1.38e+01 |
| 19 | 56 | 44.8 | 334 | 2 | ORF50 | 1.38e+01 |
| 20 | 55 | 44.8 | 383 | 2 | 1, 2-DIHYDROXYBENZYL-PY
LYTIN. | 1.38e+01 |

| | | | | | | | |
|----|----|------|------|----|--------|------------------------|----------|
| 21 | 56 | 44.8 | 416 | 2 | G92HJ3 | NADH OXIDASE (FRAGMENT | 1.38e+01 |
| 22 | 56 | 44.8 | 423 | 2 | G92HJ2 | NADH OXIDASE (FRAGMENT | 1.38e+01 |
| 23 | 56 | 44.8 | 424 | 2 | G92HJ1 | NADH OXIDASE (FRAGMENT | 1.38e+01 |
| 24 | 56 | 44.8 | 486 | 2 | G2G3P2 | OLIGOPEPTIDE TRANSPORT | 1.38e+01 |
| 25 | 56 | 44.8 | 929 | 2 | G9XJW7 | HYPOTHETICAL 101.7 KD | 1.38e+01 |
| 26 | 55 | 44.0 | 204 | 4 | G9V555 | BK150C2.3 (PUTATIVE N | 2.10e+01 |
| 27 | 55 | 44.0 | 413 | 14 | C41627 | ENVELOPE GLYCOPROTEIN | 2.10e+01 |
| 28 | 55 | 44.0 | 430 | 5 | G9XXV2 | MAPK-RELATED KINASE | 2.10e+01 |
| 29 | 55 | 44.0 | 524 | 1 | O59209 | 514AA LONG HYPOTHETICA | 2.10e+01 |
| 30 | 55 | 44.0 | 608 | 5 | O96849 | SLIME MOLD (DICTYOSTE | 2.10e+01 |
| 31 | 55 | 44.0 | 608 | 5 | 23898 | HYPOTHETICAL 70.4 KO P | 2.10e+01 |
| 32 | 55 | 44.0 | 888 | 2 | 15554 | 101 KDA PROTEIN PRECUR | 2.10e+01 |
| 33 | 55 | 44.0 | 1039 | 1 | G9F27 | 1039AA LONG HYPOTHETIC | 2.10e+01 |
| 34 | 55 | 44.0 | 1038 | 5 | C77306 | PFC0105W PROTEIN. | 2.10e+01 |
| 35 | 55 | 44.0 | 3078 | 5 | O26031 | VARIANT-SPECIFIC SUREA | 2.10e+01 |
| 36 | 55 | 44.0 | 3844 | 5 | O94646 | AARL1 PROTEIN (FRAGMEN | 2.10e+01 |
| 37 | 54 | 43.2 | 330 | 2 | G92HJ6 | HYDRATASE/ALDOLASE PHN | 3.16e+01 |
| 38 | 54 | 43.2 | 337 | 2 | G92HJ3 | HOLLIDAY JUNCTION HELI | 3.16e+01 |
| 39 | 54 | 43.2 | 1064 | 10 | C44884 | LYSINE-KETOGLUTARATE R | 3.16e+01 |
| 40 | 54 | 43.2 | 1064 | 10 | C0455 | LYSINE-KETOGLUTARATE R | 3.16e+01 |
| 41 | 54 | 43.2 | 1184 | 5 | O7752 | MALP4.27 PROTEIN | 3.16e+01 |
| 42 | 54 | 43.2 | 1200 | 5 | O92577 | PFC0860W PROTEIN | 3.16e+01 |
| 43 | 54 | 43.2 | 1308 | 5 | C96129 | PREDICTED MEMBRANE ASS | 3.16e+01 |
| 44 | 54 | 43.2 | 2708 | 5 | O15791 | CHLOROQUINE RESISTANCE | 3.16e+01 |
| 45 | 54 | 43.2 | 2815 | 5 | O15792 | STRAIN HB3 CG2 (CG2). | 3.16e+01 |

ALL NMENTS

| | | | |
|---|-----------------------------------|-------------------------------|-----------|
| RESULT 1 | PRELIMINARY: | PRT: | 375 AA |
| ID Q96395 | | | |
| AC Q96395 | | | |
| DI 01-FEB-1997 (TRENBLrel. 02, Created) | | | |
| DI 01-FEB-1997 (TRENBLrel. 02, Last sequence update) | | | |
| DI 01-NOV-1999 (TRENBLrel. 12, Last annotation update) | | | |
| DE CHA01 PRECURSOR. | | | |
| OS Chamaecyparis obtusa. | | | |
| OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; | | | |
| OC euphyllophytes; Spermatophyta; Coniferopsida; Coniferales; | | | |
| OC Taxodiaceae; Chamaecyparis. | | | |
| RN [1] | | | |
| RP SEQUENCE FROM N.A. | | | |
| RC TISSUE-POLLEN: | | | |
| RX MEDLINE: 96265194. | | | |
| RA SUZUKI M., KOMIYAMA N., ITOH M., ITOH H., SONE T., KINUKA K., IARAKI I., | | | |
| RA CHIA N. | | | |
| RT *Purification, characterization and molecular cloning of Cha 01, a | | | |
| RT major allergen of Chamaecyparis obtusa (Japanese cypress) pollen. | | | |
| RL Mol Immunol: 33:451-460(1996). | | | |
| DR EMBL: D45464; BAA08246.1; ... | | | |
| DR MENDEL: 7626; ChaoB:1088;7626. | | | |
| DR PFAM: PF00544; pec_lyase: 1 | | | |
| DR PRINTS: PRO0807; AMBALLERGEN. | | | |
| KW Signal. | | | |
| FT SIGNAL | 1 21 | POTENTIAL. | |
| FT CHAIN | 22 375 | CHAO1. | |
| SQ SEQUENCE | 375 AA: 40258 MW: A0981492 CRC32: | | |
| Query Match | 100.0% | Score 125; DB 10; Length 375; | |
| Best Local Similarity | 100.0% | Pred. No. 2.25e-15; | |
| Matches | 20: Conservative | 0: Mismatches | 0: Indels |
| | 0: Gaps | 0: | |
| Db 332 SSGKNEGNIYNNNEAFKVE 351 | | | |
| QY 1 SSGKNEGNIYNNNEAFKVE 20 | | | |
| | | | |
| RESULT 2 | PRELIMINARY: | PRT: | 367 AA |
| ID Q92NU7 | | | |
| AC Q92NU7 | | | |
| DI 01-MAY-1999 (TRENBLrel. 10, Created) | | | |
| DI 01-MAY-1999 (TRENBLrel. 10, Last sequence update) | | | |
| DI 01-NOV-1999 (TRENBLrel. 12, Last annotation update) | | | |

```

DB  POLLEN MAJOR ALLERGEN 1-2.
OS  Juniperus ashei (Ozark white cedar).
OC  Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC  euphyllophytes; Spermatophyta; Coniferopsida; Coniferales;
OC  Taxodiaceae; Juniperus.
RN  [1]
RP  SEQUENCE FROM N.A.
RA  MIDORO-HORIUTI T.M., GOLDBLUM R.M., KUROSKY A., WOOD T.G.,
RA  BROOKS E.G.
RI  "Molecular cloning of mountain cedar (Juniperus ashei) pollen major
RI  allergen, Jun a 1."
RC  Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR  EMBL: AF106663; AAC03609;
DR  EMBL: AF106662; AAC03608;
DR  MENDEL: 36544; Jcudas:1086;36544.
DR  MENDEL: 36545; Jcudas:1088;36545.
SQ  SEQUENCE 367 AA: 39824 MW: 4C2D8630 CRC32:

Query Match      86.4%  Score 108: DB 10: Length 367:
Best Local Similarity 85.0%  Pred. No. 4.37e-11:
Matches 17: Conservative 1: Mismatches 2: Indels 0: Gaps 0:

DB  332 SSGKTEENIYNSNEAKVE 35:
QY  1 SSGKTEENIYNNNEAKVE 20

RESULT
ID  Q9ZSS7 PRELIMINARY: PRT: 1198 AA.
AC  Q9ZSS7
DI  01-MAY-1999 (TrEMBLrel: 10, Created)
DI  01-MAY-1999 (TrEMBLrel: 10, Last sequence update)
DI  01-MAY-1999 (TrEMBLrel: 10, Last annotation update)
DE  DYNEIN HEAVY CHAIN ISOFORM DHC1B (FRAGMENT).
OS  DHC1B.
OC  Chlamydomonas reinhardtii.
OC  Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC  Chlamydomonadales; Chlamydomonas.
RN  [1]
RP  SEQUENCE FROM N.A.
RA  PAZOUK G.J., DICKERT B.L., WITMAN G.B.
RI  "The dynein heavy chain isoform DHC1B is required for flagellar
RI  assembly."
RC  Submitted (OCT 1998) to the EMBL/GenBank/DBJ databases.
DR  EMBL: AF096277; AAC99457;1;
DR  NON_TER 1
FT  NON_TER 1198
SQ  SEQUENCE 1198 AA: 134914 MW: 45558276 CRC32:

Query Match      55.2%  Score 69: DB 10: Length 1198:
Best Local Similarity 35.3%  Pred. No. 4.01e-02:
Matches 6: Conservative 10: Mismatches 1: Indels 0: Gaps 0:

DB  159 REESAKYKEAEAKID 175
QY  4 KNEGTNIYNNNEAKVE 20

RESULT
ID  Q08144 PRELIMINARY: PRT: 397 AA.
AC  Q08144
DI  01-NOV-1996 (TrEMBLrel: 01, Created)
DI  01-NOV-1996 (TrEMBLrel: 01, Last sequence update)
DI  01-NOV-1998 (TrEMBLrel: 08, Last annotation update)
DE  CHROMOSOME XV READING FRAME ORF Y01018C.
OS  Saccharomyces cerevisiae (Baker's yeast).
OC  Eukaryota; Fungi; Ascomycota; Hemiascomycetes; Saccharomycetales;
OC  Saccharomycetaceae; Saccharomyces.
RN  [1]
RP  SEQUENCE FROM N.A.
RA  HUGHES B., POHL T.M.
RI  Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN  [2]

RP  SEQUENCE FROM N.A.
RA  MIPS:
RA  Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR  EMBL: 274760; CAA95017;1;
DR  PFAM: PF00894; Syntaxin1.1
SQ  SEQUENCE 397 AA: 45875 MW: 79FB1A99 CRC32:

Query Match      50.4%  Score 63: DB 3: Length 397:
Best Local Similarity 51.3%  Pred. No. 6.52e-01:
Matches 8: Conservative 6: Mismatches 1: Indels 0: Gaps 0:

DB  342 SSGRNGSNKYNND 356
QY  1 SSGRNGSNKYNNE 15

RESULT
ID  Q9ZHT7 PRELIMINARY: PRT: 420 AA.
AC  Q9ZHT7
DI  01-MAY-1999 (TrEMBLrel: 10, Created)
DI  01-MAY-1999 (TrEMBLrel: 10, Last sequence update)
DI  01-NOV-1999 (TrEMBLrel: 12, Last annotation update)
DE  NADH OXIDASE (FRAGMENT).
OS  Serpulina sp. P280/1.
OC  Bacteria; Spirochaetales; Spirochaetaceae; Serpulina.
RN  [1]
RP  SEQUENCE FROM N.A.
RA  ATYEO R.F., JENSEN N.S., STANTON T.B., HAMPTON D.J.
RI  Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR  EMBL: AF060815; AAC78824;1;
DR  HSP: P37062; 1JOA.
FT  NON_TER 1
FT  NON_TER 420
SQ  SEQUENCE 420 AA: 46594 MW: 1E88D66 CRC32:

Query Match      48.8%  Score 61: DB 2: Length 420:
Best Local Similarity 35.3%  Pred. No. 1.60e-00:
Matches 6: Conservative 6: Mismatches 6: Indels 0: Gaps 0:

DB  54 KGEIDVYMGHEVWKID 70
QY  4 KNEGTNIYNNNEAKVE 20

RESULT
ID  Q73616 PRELIMINARY: PRT: 149 AA.
AC  Q73616
DI  01-AUG-1998 (TrEMBLrel: 07, Created)
DI  01-AUG-1998 (TrEMBLrel: 07, Last sequence update)
DI  01-AUG-1998 (TrEMBLrel: 07, Last annotation update)
DE  TRANSKETOLASE (FRAGMENT).
OS  IK.
OC  Xenopus laevis (African clawed frog).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Amphibia.
OC  Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae; Xenopodinae;
OC  Xenopus.
RN  [1]
RP  SEQUENCE FROM N.A.
RA  SPINDLER B., VERREY F.
RI  Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
DR  EMBL: AJ224484; CAA11971;1;
FT  NON_TER 1
FT  NON_TER 126
SQ  SEQUENCE 126 AA: 13539 MW: B740600F CRC32:

Query Match      48.0%  Score 60: DB 13: Length 126:
Best Local Similarity 33.3%  Pred. No. 2.48e-00:
Matches 6: Conservative 7: Mismatches 5: Indels 0: Gaps 0:

DB  21 TSPEDAVIYSSTEEFKI 38
QY  1 TSPEDAVIYSSTEEFKI 38

```

QY 2 SSKNEGTNNNEAFKV 19

RESULT 7
ID Q45841 PRELIMINARY: PRT: 146 AA.
AC Q45841:
DI 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DE 01-NOV-1996 (TRENBLrel. 07, Last annotation update)
CE 17 KB REMAGGLUTININ COMPONENT.
GN HEM17/B.
GS Clostridium botulinum.
OC Bacteria: Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-LAMANA:
RA YANG G., RHEE S., JUNG H., YANG K.
RL Submitted (APR-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: U2443; AAA9056.1; .
SQ SEQUENCE 146 AA: 15908 MW: 6150A4C9 CRC32:

Query Match 47.2% Score 59: DB 2: Length 146;
Best Local Similarity 58.3%; Pred. No. 3.84e+00;
Matches 7: Conservative 4: Mismatches 1: Indels 0: Gaps 0:
Db 133 DIVNSQMKLE 144
QY 9 NIYNNNEAFKVE 20

RESULT 8
ID Q45878 PRELIMINARY: PRT: 146 AA.
AC Q45878:
DI 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DE HA17.
GN HA17 OR HA-17.
OC Clostridium botulinum.
OC Bacteria: Firmicutes; Bacillus/Clostridium group; Clostridiaceae;
OC Clostridium.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 96291660.
RA HENDERSON I., WHELAN S.M., DAVIS T.J., MINTON N.P.
RL "Genetic characterisation of the botulinum toxin complex of
Clostridium botulinum strain NGC 2916."
RT FEMS Microbiol. Lett. 140:151-156(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN-CDC 328; (ATCC 43757);
RA MEDLINE: 98440323.
RA SANTOS-BUELGA J., COLLINS M.D., EAST A.K.
RL "Characterization of the genes encoding the botulinum neurotoxin
complex in a strain of Clostridium botulinum producing type B and F
neurotoxins."
RI Curr. Microbiol. 37:312-318(1998).
RL Curr. Microbiol. 37:312-318(1998).
DR EMBL: L42537; AAB42188.1; .
DR EMBL: Y13630; CAA73964.1; .
SQ SEQUENCE 146 AA: 17034 MW: C190B1B CRC32:

Query Match 47.2% Score 59: DB 2: Length 146;
Best Local Similarity 58.3%; Pred. No. 3.84e+00;
Matches 7: Conservative 4: Mismatches 1: Indels 0: Gaps 0:
Db 133 DIVNSQMKLE 144
QY 9 NIYNNNEAFKVE 20
RESULT 9
ID O83926 PRELIMINARY: PRT: 254 AA.

O83926:
AC DT 01-NOV-1998 (TRENBLrel. 08, Created)
DI 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TRENBLrel. 10, Last annotation update)
DE FLAGELLAR BASAL-BODY ROD PRO. FIN (FLGG-1).
GN TPC960.
OC Treponema pallidum.
OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 98332770.
RA FRASER C.M., NORRIS S.J., WEINSTOCK G.M., WHITE O., SUTTON G.G.
RA DODSON R., GWINN M., HICKEY E.K., CLAYTON R., KETCHUM K.A.,
RA SODERREN E., HARDHAM J.M., MCLEOD M.P., SALZBERG S., PETERSON J.,
RA KHALAK H., RICHARDSON D., HOWELL J.K., CHIDAMBARAM M., UTTERBACK T.,
RA McDONALD L., ARTIACH P., BOWMAN C., COTTON M.D., FUJII C., GARLAND S.,
RA HATCH B., HORST K., ROBERTS K., WATTHEY L., WEIDMAN J., SMITH H.O.,
RA VENTER J.C.
RL "Complete genome sequence of Treponema pallidum, the syphilis
spirochete."
RL Science 281:375-388(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA FRASER C.M., NORRIS S.J., WEINSTOCK G.M., WHITE O., SUTTON G.G.
RA DODSON R., GWINN M., HICKEY E.K., CLAYTON R., KETCHUM K.A.,
RA SODERREN E., HARDHAM J.M., MCLEOD M.P., SALZBERG S., PETERSON J.,
RA KHALAK H., RICHARDSON D., HOWELL J.K., CHIDAMBARAM M., UTTERBACK T.,
RA McDONALD L., ARTIACH P., BOWMAN C., COTTON M.D., FUJII C., GARLAND S.,
RA HATCH B., HORST K., ROBERTS K., WATTHEY L., WEIDMAN J., SMITH H.O.,
RA VENTER J.C.
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AE001264; AAC65615.1;
DR TIGR: IP0560;
DR PFAM: PF00460; flg_bb_rnd; .
KW Flagella.
SQ SEQUENCE 264 AA: 26597 MW: FC2A7F39 CRC32:
Query Match 47.2% Score 59: DB 2: Length 264;
Best Local Similarity 40.0%; Pred. No. 3.84e+00;
Matches 6: Conservative 4: Mismatches 5: Indels 0: Gaps 0:
Db 109 DGIYAYTROGSKVD 123
QY 6 EGTNNNEAFKVE 20
RESULT 10
ID Q19059 PRELIMINARY: PRT: 246 AA.
AC Q19059:
DI 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DE 01-NOV-1999 (TRENBLrel. 12, Last annotation update)
GN SIMILAR TO HYDRATASE-DEHYDROGENASE-EPIMERASE.
OC E04F6.3.
OC Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae;
OC Rhabditia; Rhabditidae; Rhabditidae; Pelodidae; Caenorhabditis.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-BRISTOL N2.
RA MEDLINE: 94150718.
RA WILSON R., AINSKOUGH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA CRAXTON M., DEAR S., DU Z., DORBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAWKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIFKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN K., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.,
RL "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans."

RL Nature 368:32-38(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2:
 RA PAULEY A.:
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BRISTOL N2:
 RA WATERSTON R.:
 RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U28543; AAA68358.1;
 DR PFAM: PF01575; Maoc_like; 2;
 SQ SEQUENCE 298 AA: 32721 MW: 556488 FE CRC32:

Query Match 47.2% Score 59 DB 5: Length 298;
 Best Local Similarity 36.8% Pred. No. 3,84e-00;
 Matches 7: Conservative 7: Mismatches 5: Indels 0: Gaps 0:

DB 32 ARAKEDLCYVYVNHEDFKV 50
 QY 1 SSQKNEGINVYNNNEAFK 19

RESULT 11
 ID Q59078 PRELIMINARY: PRT: 290 AA.
 AC Q59078;
 DT 01-NOV-1996 (TRENBLrel. 01: Created)
 DI 01-NOV-1996 (TRENBLrel. 01: Last sequence update)
 DE BETA-AGARASE:1 PRECURSOR (EC 3.2.2.81).
 US Alteromonas atlantica (Pseudomonas atlantica).
 OC Bacteria; Proteobacteria; gamma subdivision; Alteromonadaceae;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA CUAN C., RIGGS P.D., BENNER J.S., ZHU P., CHOU D., HALL I.S.,
 RA KEATING C.P.:
 RC Submitted (MAP 1996) to the EMBL/GenBank/DBJ databases.
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF 1,4-BETA-D-GALACTOSIDIC LINKAGES
 CC IN AGAROSE, GIVING THE TETRAMER AS THE PREDOMINANT PRODUCT.
 DR EMBL: M73783; AAA91688.1;
 KW Signal: Hydrophobic; Glycosylase;
 ET CHAIN 1 21 290
 SQ SEQUENCE 290 AA: 32795 MW: 7550180 FE CRC32:

Query Match 46.4% Score 58 DB 2: Length 290;
 Best Local Similarity 40.0% Pred. No. 5,92e-00;
 Matches 5: Conservative 5: Mismatches 3: Indels 0: Gaps 0:

DB 187 KDAGSNVYNGEYR 201
 QY 4 KNEGTINYNNNEAFK 18

RESULT 12
 ID Q9YW42 PRELIMINARY: PRT: 379 AA.
 AC Q9YW42;
 DT 01-MAY-1999 (TRENBLrel. 10: Created)
 DI 01-MAY-1999 (TRENBLrel. 10: Last sequence update)
 DE ORF MSV050 HYPOTHETICAL PROTEIN.
 GN MSV050

OC Melanoplus sanguinipes entomopoxvirus.
 OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Entomopoxvirinae;
 OC Entomopoxvirus B.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TUCSON:
 RX MEDLINE: 99102612.
 RA AFONSO C.L., TULMAN E.R., LU Z., OMA E., KUTISH G.F., ROCK D.L.:
 R: *The genome of Melanoplus sanguinipes entomopoxvirus.*;

RL J. Virol. 73:533-552(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TUCSON:
 RA AFONSO C.L., TULMAN E.R., LU Z., OMA E., KUTISH G.F., ROCK D.L.:
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF063866; AAC97619.1;
 SQ SEQUENCE 379 AA: 43956 MW: 72541767 FE CRC32:

Query Match 46.4% Score 58 DB 14: Length 379;
 Best Local Similarity 61.5% Pred. No. 5,92e-00;
 Matches 8: Conservative 3: Mismatches 2: Indels 0: Gaps 0:

DB 201 GTNGYNNESFKI 213
 QY 7 GINIYNNNEAFK 19

RESULT 13
 ID Q87980 PRELIMINARY: PRT: 412 AA.
 AC Q87980;
 DT 01-NOV-1998 (TRENBLrel. 08: Created)
 DI 01-NOV-1998 (TRENBLrel. 08: Last sequence update)
 DE 01-NOV-1998 (TRENBLrel. 08: Last annotation update)
 DE HYPOTHETICAL 50.7 KD PROTEIN.
 GN BBLP51.07.
 OS Bordetella bronchiseptica.
 OC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
 OC Bordetella.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CN7635E;
 RA STEVENS K., CHURCHER C.M., BADOCK K.L.:
 RC Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CN7635E;
 RA PARHILL J., PRESTON A., MASKELL D.J., FARRELL B.G.:
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: A207747; CAA07646.1;
 KW Hypothetical; protein.
 SQ SEQUENCE 412 AA: 50744 MW: 4789483 FE CRC32:

Query Match 45.4% Score 58 DB 2: Length 412;
 Best Local Similarity 40.0% Pred. No. 5,92e-00;
 Matches 6: Conservative 5: Mismatches 4: Indels 0: Gaps 0:

DB 73 MBESSQYVYRGEDFR 87
 QY 4 KNEGTINYNNNEAFK 18

RESULT 14
 ID O96205 PRELIMINARY: PRT: 2206 AA.
 AC O96205;
 DT 01-MAY-1999 (TRENBLrel. 10: Created)
 DI 01-MAY-1999 (TRENBLrel. 10: Last sequence update)
 DE 01-MAY-1999 (TRENBLrel. 10: Last annotation update)
 DE HYPOTHETICAL 264.1 KD PROTEIN.
 GN PFB0560K.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 99021743.
 RA GARDNER M.J., TETTELIN H., CARUCCI D.J., CUMMINGS L.M., ARAVIND L.,
 RA KOONIN E.V., SHALLOO S., MASON T., YU K., FUJII C., PEDERSON J.,
 RA SHEN K., JING J., ASTON C., LAI Z., SCHWARTZ D.C., PERTEA M.,
 RA SALZBERG S., ZHOU L., SUTTON G.G., CLAYTON R., WHITE C., SMITH H.G.,
 RA FRASER C.M., ADAMS M.D., VENTER J.C., HOFFMAN S.L.:
 RI Chromosome 2 sequence of the human malaria parasite Plasmodium
 R: falciparum.*;
 RL Science 282:1126-1132(1998).

DP EMBL: AEO01403; AAC71901.1; ...
 KW Hypothetical protein.
 SQ SEQUENCE 2206 AA: 264100 MW: AEE13480 CRC32:

 Query Match 45.6%; Score 57; DB 5; Length 2206;
 Best Local Similarity 35.7%; Pred. No. 9 07e+00;
 Matches 5; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

 DB 353 NNEIDVYNDKDNF 376
 QY 4 KNEGTNIYNNNEAF 17

 RESULT 15
 ID 077336 PRELIMINARY: PRI: 4550 AA.
 AC 077336:
 DI 01-NOV-1998 (TRENBLrel. 08, Created)
 DI 01-NOV-1998 (TRENBLrel. 08, Last sequence update)
 DI 01-NOV-1999 (TRENBLrel. 12, Last annotation update)
 DE PFC0425W PROTEIN.
 GN PFC0425W.
 OS Plasmodium falciparum.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MUNGALL K., LAWSON D., BARRELL B.:
 RC Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: 798547; CAB11211.1;
 SQ SEQUENCE 4550 AA: 534109 MW: CC698096 CRC32:

 Query Match 45.6%; Score 57; DB 5; Length 4550;
 Best Local Similarity 18.8%; Pred. No. 9 07e+00;
 Matches 3; Conservative 9; Mismatches 4; Indels 0; Gaps 0;

 DB 831 EDDDDIYDDDDNFED 846
 QY 5 NEGNIYNNNEAFKE 20

Search completed: Mon Jun 19 16:15:11 2000
 Job time : 12 secs.

W B E R E F (TM)

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MParch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Mon Jun 19 16:17:34 2000; MasPar time 3.82 seconds
Tabular output not generated.
80,607 Million cell updates/sec

Title: >US-09-142-524A-13
Description: (1-13) from US09142524A.pap
Perfect Score: 94
Sequence: VFIRKVSNIIVHG 13

Scoring table: PAM 150
Gap 15

Searched: 186963 seqs, 23686106 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-sequences35
:igoneseqp

Statistics: Mean 18.336; Variance 56.950; scale 0.322

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Length | Seq. ID | Description | Pred. No. |
|------------|-------|--------------|---------|-------------------------------|-----------|
| 1 | 94 | 100.0 | 1 | R89291 Japanese cedar pollen | 4.88e-03 |
| 2 | 94 | 100.0 | 1 | Cry j 1 pollen allergen | 2.88e-03 |
| 3 | 94 | 100.0 | 1 | R82501 Cry j 1 Japanese Cedar | 2.88e-03 |
| 4 | 94 | 100.0 | 30 | T-cell epitope peptide | 2.88e-03 |
| 5 | 94 | 100.0 | 30 | Multi-epitope peptide | 2.88e-03 |
| 6 | 94 | 100.0 | 80 | Multi-epitope peptide | 2.88e-03 |
| 7 | 94 | 100.0 | 195 | Multi-epitope peptide | 2.88e-03 |
| 8 | 94 | 100.0 | 134 | Multi-epitope peptide | 2.88e-03 |
| 9 | 94 | 100.0 | 353 | Japanese cedar pollen | 2.88e-03 |
| 10 | 94 | 100.0 | 374 | Cedar pollen allergen | 2.88e-03 |
| 11 | 94 | 100.0 | 374 | Cry j 1 | 2.88e-03 |
| 12 | 94 | 100.0 | 374 | Cry j 1 Japanese Cedar | 2.88e-03 |
| 13 | 94 | 100.0 | 374 | Cry j 1 pollen allergen | 2.88e-03 |
| 14 | 82 | 87.2 | 15 | R60166 Japanese cedar pollen | 2.88e-03 |
| 15 | 64 | 73.4 | 20 | Residues 106-120 of Cr | 7.22e-02 |
| 16 | 59 | 73.4 | 20 | Cry j 1 pollen allergen | 2.12e-02 |
| 17 | 57 | 71.3 | 370 | Cry j 1 Japanese Cedar | 2.12e-02 |
| 18 | 54 | 68.1 | 9 | Residues 109-117 of Cr | 3.51e-02 |
| 19 | 54 | 68.1 | 31 | Multi-epitope peptide | 7.46e-02 |
| 20 | 54 | 68.1 | 31 | Multi-epitope peptide | 7.46e-02 |
| 21 | 54 | 68.1 | 20 | T-cell epitope peptide | 1.57e-02 |
| 22 | 54 | 64.9 | 354 | Chamaecyparis obtusa p | 1.57e-02 |
| 23 | 54 | 64.9 | 1 | Japanese cypress pol.e | 1.57e-02 |

| | | | | | | | |
|----|----|------|------|---|--------|-------------------------|----------|
| 24 | 61 | 64.9 | 357 | 1 | R45577 | Jun s I. | 1.57e-02 |
| 25 | 61 | 64.9 | 375 | 1 | W04345 | Chamaecyparis obtusa p | 1.57e-02 |
| 26 | 56 | 59.6 | 292 | 1 | R04895 | Penicillium-Insulin | 5.30e-01 |
| 27 | 56 | 59.6 | 408 | 1 | W94245 | A. orientalis glycosyl | 5.30e-01 |
| 28 | 56 | 59.6 | 408 | 1 | W94245 | A. orientalis glycosyl | 5.30e-01 |
| 29 | 54 | 57.4 | 367 | 1 | W56279 | Babesia microti BMN1-3 | 8.54e-01 |
| 30 | 54 | 57.4 | 367 | 1 | W56301 | Babesia microti antiige | 8.54e-01 |
| 31 | 52 | 55.3 | 192 | 1 | R89533 | Hepatitis C virus iscl | 1.37e-02 |
| 32 | 52 | 55.3 | 192 | 1 | R89662 | Hepatitis C virus enve | 1.37e-02 |
| 33 | 51 | 54.3 | 40 | 1 | R70757 | Spider venom FIL-376 N | 1.73e-02 |
| 34 | 51 | 54.3 | 308 | 1 | W32430 | Mycobacterium tubercul | 1.73e-02 |
| 35 | 51 | 54.3 | 309 | 1 | W32362 | Mycobacterium tubercul | 1.73e-02 |
| 36 | 51 | 54.3 | 309 | 1 | W64302 | M. tuberculosis immuno | 1.73e-02 |
| 37 | 51 | 54.3 | 309 | 1 | W61665 | M. tuberculosis immuno | 1.73e-02 |
| 38 | 51 | 54.3 | 857 | 1 | W67702 | Staphylococcus aureus | 1.73e-02 |
| 39 | 50 | 53.2 | 316 | 1 | W37356 | Psst (WbpI) protein in | 2.18e-02 |
| 40 | 50 | 53.2 | 552 | 1 | W16135 | Omega-cychohexane fatt | 2.18e-02 |
| 41 | 50 | 53.2 | 672 | 1 | W96657 | Staphylococcus aureus | 2.18e-02 |
| 42 | 50 | 53.2 | 866 | 1 | W96556 | Staphylococcus aureus | 2.18e-02 |
| 43 | 50 | 53.2 | 908 | 1 | R33429 | E. coli heat shock pro | 2.18e-02 |
| 44 | 49 | 52.1 | 1565 | 1 | W18305 | Photobacterium luminesc | 2.74e-02 |
| 45 | 49 | 52.1 | 1565 | 1 | W56566 | Toxin TccB, encoded by | 2.74e-02 |

ALIGNMENTS

RESULT 1
ID R89291 standard; peptide: 15 AA.
AC R89291:
DE 12-MAR-1995 (first entry)
DI Japanese cedar pollen allergen Cryj I derived T-cell epitope peptide.
KW Japanese cedar; pollen allergen; Cryj I; T-cell epitope; peptides.
KW prevention; treatment; cryptomeria pollenosis.
OS Cryptomeria japonica.
PN J07118295-A.
PD 09-MAY-1995.
PF 20-OCT-1993; 262626.
PR 20-OCT-1993; JP-262626.
PA (MEIP) MEIJI MILK PROD CO LI.
DR WPI: 95-203834/27.
PI New cryptomeria pollen allergen T-cell epitope peptide - used for:
PI prevention, treatment and investigation of Japanese cedar pollenosis
PS Claim 5; page 2; 8pp: Japanese.
CC R75388 is the Japanese cedar pollen allergen Cryj I, from which the
CC T-cell epitope peptides R89289-R89295 were derived. The peptides
CC can be used for the prevention and treatment of cryptomeria
CC pollenosis, and also for the investigation of pollenosis.
SQ Sequence 15 AA:

Query Match: 100.0%; Score 94; DB 1; Length 15.
Best Local Similarity 100.0%; Pred. No. 2.88e-03.
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 3 VFIRKVSNIIVHG 15
QY 1 VFIRKVSNIIVHG 13
|||||

RESULT 2
ID R45552 standard; Protein: 20 AA.
AC R45552:
DI 13-JUL-1994 (first entry)
DE Cry j 1 pollen allergen peptide CJI-11.
KW Japanese cedar; detection; allergy; treatment; diagnosis;
KW T cell epitope; sensitivity.
OS Cryptomeria japonica.
PN W09601560-A.
PD 20-JAN-1994.
PF 15-JAN-1993; 000139.
PR 10-JUL-1992; WC-005661.
PR 01-SEP-1992; JS-938990.
PA (IMMU-) IMMUNOLOGIC PHARM CORP.
PI Bond Jf. Garraan RD. Griffith. J. Kuo M. Pollock J.

DA WPI: 94-03566/24.
 PF Antigens derived from Japanese cedar pollen allergen Cry j I -
 PA contain at least two T cell epitope(s), used to treat or diagnose
 PT allergy.
 PS Claim 1: Fig 13: 137pp: English.
 CC The sequence is that of an isolated peptide of the Japanese cedar
 pollen allergen Cry j I (amino acids 101-120). The peptide, CJI-11,
 can be used for the treatment and diagnosis of allergies associated
 with Japanese cedar pollen. It has enhanced therapeutic properties
 but reduced side effects compared to naturally occurring allergens.
 SQ Sequence 20 AA:

Query Match 100.0%; Score 94; DB 1; Length 20;

Best Local Similarity 100.0%; Pred No. 2,89e-03;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 8 VFIRKRVSNVIING 20

QY 1 VFIRKRVSNVIING 13

RESULT 3

ID R92501 standard; Protein: 20 AA

AC R92501;

DE Cry j I Japanese Cedar pollen allergen peptide fragment (CJI-11).

KW Cry j I: Japanese cedar pollen allergen; modified: drug production;

KW Allergy; Cryptomeria japonica.

OS Cryptomeria japonica.

PN W0952786-A1.

PR 19-OCT-1995.

PF 06-APR-1995; U04249.

PR 08-APR-1994; US-226248.

PR 06-DEC-1994; US-150225.

PA (IMMUNO) IMMUNOLOGIC PHARM CORP.

PI Chen X., Evans S., Fahren HM., Kuo X., Powers SP.

PT Shaked Z.

PS WPI: 95-36630/47.

CC Modified Cryptomeria japonica (Cry j) I peptide(s) - useful for

treating allergy to Japanese cedar pollen allergen or

immunologically cross reactive allergens

IS Disclosure figure 2; 60pp: English.

CC Novel peptides of Cry j I have been modified as a part of a

preformulation scheme to develop an optimised drug product for

therapeutic treatment of humans suffering from allergy to Japanese

cedar pollen allergen or an allergen which is immunologically cross

reactive with Japanese cedar pollen allergen. Such modified peptides

possess certain characteristics which render them particularly

suitable for drug product formulation. Peptide fragments of Cry j I

modified and re-modified, are given in R9249-182925. This peptide

fragment corresponds to amino acids 101-120 of the allergen mature

protein

SQ Sequence 20 AA:

Query Match 100.0%; Score 94; DB 1; Length 20;

Best Local Similarity 100.0%; Pred No. 2,89e-03;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 6 VFIRKRVSNVIING 20

QY 1 VFIRKRVSNVIING 13

RESULT 4

ID W44683 standard; peptide: 20 AA.

AC W44683;

DE 01-MAY-1998 (first entry)

CC T-cell epitope peptide #2 of sugi pollen antigen.

KW T-cell epitope, sugi pollen antigen; sug; pollinosis.

OS Synthetic.

PN J1000782-A

PR 13-JAN-1998

PF WPI: 94-03566/24.

PR 24-JUN-1996; JP-163287.

PA (DAIL) DAICEL CHEM IND LTD.

PA (MEIJ) MEIJI SEIKA KAISHA LTD.

DR WPI: 98-133630/13.

PT T-cell epitope peptide of sugi pollen antigen - useful in the

treatment of sugi pollinosis

PS Claim 1: Page 4: 14pp: Japanese.

CC T-cell epitope peptides W44682-83 and their derivatives react with

sugi pollinosis patient peripheral blood T lymphocytes. A composition

prepared by combining at least 2 of the above peptides and/or their

CC derivatives is used for the prevention and treatment of sugi pollinosis.

SQ Sequence 30 AA:

Query Match 100.0%; Score 94; DB 1; Length 30;

Best Local Similarity 100.0%; Pred No. 2,89e-03;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 8 VFIRKRVSNVIING 20

QY 1 VFIRKRVSNVIING 13

RESULT 5

ID W27369 standard; peptide: 80 AA.

AC W27369;

DE 24-MAR-1998 (first entry)

CC Multi-epitope peptide used as immunotherapeutic agent #1.

KW Multi-epitope peptide; immunotherapeutic agent; allergic disease;

KW T-cell epitope region; allergen; lymphocyte; immunoglobulin E.

OS Synthetic.

PN W09732600-A1.

PR 12-SEP-1997.

PR 10-MAR-1997; J00740.

PR 10-MAR-1996; JP-080702.

PA (MEIJ) MEIJI MILK PROD CO LTD.

PI Cairiki K., Iwara A., Kino K., Kume A., Sone T.

PS WPI: 97-470495/43.

CC Peptide is immunotherapeutic agent to treat allergic diseases.

PT contains multi-epitope peptide containing T cell epitope region

from different allergens

PS Claim 6: Page 31: 58pp: Japan.

CC The present sequence represents a multi-epitope peptide which is used as

a new immunotherapeutic agent. It comprises T cell epitope regions from 2

or more different allergens (preferably linked via arginine or lysine

dimers), where the T cell epitope regions have a positivity index

greater than 100 as measured in a patient group responding to the

CC allergen; have at least 70% reactivity with lymphocytes from patients

CC responding to the allergen; and are not reactive with immunoglobulin E

CC (IgE) antibodies from patients responsive to the allergen. The agent can

CC be used to prevent and treat a wide variety of allergic diseases caused by

CC desensitisation. Side effects and those mediated by IgE are reduced.

SQ Sequence 80 AA:

Query Match 100.0%; Score 94; DB 1; Length 80;

Best Local Similarity 100.0%; Pred No. 2,89e-03;

Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 16 VFIRKRVSNVIING 28

QY 1 VFIRKRVSNVIING 13

RESULT 6

ID W27370 standard; peptide: 105 AA.

AC W27370;

DE 24-MAR-1998 (first entry)

CC Multi-epitope peptide used as immunotherapeutic agent #2.

KW Multi-epitope peptide; immunotherapeutic agent; allergic disease;

KW T-cell epitope region; allergen; lymphocyte; immunoglobulin E.

OS Synthetic.

PN W09732600-A1.

PR 12-SEP-1997.

PF 10-MAR-1997: J00740.
PR 10-MAR-1996: JP-080702.
PA (MEIP) MEIJI MILK PROD CO LTD.
PI Dairiki K, Iwama A, Kino K, Kume A, Sone T;
PW: 97-470495/43.
BI Peptide immunotherapeutic agent to treat allergic diseases -
PI contains multi-epitope peptide containing T cell epitope regions
PS Claim 6: Page 31: 58pp; Japanese.
CC The present sequence represents a multi-epitope peptide which is used as
CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2
CC or more different allergens (preferably linked via arginine or lysine
CC dimers), where the T cell epitope regions have a positivity index
CC greater than 100 as measured in a patient group responding to the
CC allergen; have at least 70% reactivity with lymphocytes from patients
CC responding to the allergen; and are not reactive with immunoglobulin E
CC (IgE) antibodies from patients responsive to the allergen. The agent can
CC be used to prevent and treat a wide variety of allergic diseases, e.g. by
CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
SV Sequence 105 AA:
Query Match 100.0%; Score 94; DB 1; Length 105;
Best Local Similarity 100.0%; Pred. No. 2.88e-03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 16 VFIKRVSNVILHG 28
QY 1 VFIKRVSNVILHG 13
RESULT 7
ID W27371 standard; peptide; 134 AA.
AC W27371 (first entry)
DE Multi-epitope peptide used as immunotherapeutic agent #3.
KW Multi-epitope peptide; immunotherapeutic agent; allergic disease;
KW T-cell epitope region; allergen; lymphocyte; immunoglobulin E.
CS Synthetic.
PN W09732600-A1.
PC 12-SEP-1997.
PF 10-MAR-1997: J00740.
PR 10-MAR-1996: JP-080702.
PA (MEIP) MEIJI MILK PROD CO LTD.
PI Dairiki K, Iwama A, Kino K, Kume A, Sone T;
PW: 97-470495/43.
BI Peptide immunotherapeutic agent to treat allergic diseases -
PI contains multi-epitope peptide containing T cell epitope regions
PS Claim 6: Page 31: 58pp; Japanese.
CC The present sequence represents a multi-epitope peptide which is used as
CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2
CC or more different allergens (preferably linked via arginine or lysine
CC dimers), where the T cell epitope regions have a positivity index
CC greater than 100 as measured in a patient group responding to the
CC allergen; have at least 70% reactivity with lymphocytes from patients
CC responding to the allergen; and are not reactive with immunoglobulin E
CC (IgE) antibodies from patients responsive to the allergen. The agent can
CC be used to prevent and treat a wide variety of allergic diseases, e.g. by
CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
SV Sequence 134 AA:
Query Match 100.0%; Score 94; DB 1; Length 134;
Best Local Similarity 100.0%; Pred. No. 2.88e-03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 16 VFIKRVSNVILHG 28
QY 1 VFIKRVSNVILHG 13
RESULT 8
ID W75388 standard; protein; 353 AA.
AC W75388

DT 12-MAR-1996 (first entry)
DE Japanese cedar pollen allergen Cryj I.
KW Japanese cedar; pollen allergen; Cryj I; T-cell epitope; peptides;
KW prevention; treatment; cryptomeria pollinosis.
OS Cryptomeria japonica.
FH Key Location/Qualifiers
FT peptide 51..75 /note= "T-cell epitope peptide"
FT peptide 91..105 /note= "T-cell epitope peptide"
FT peptide 106..120 /note= "T-cell epitope peptide"
FT peptide 146..160 /note= "T-cell epitope peptide"
FT peptide 211..225 /note= "T-cell epitope peptide"
FT peptide 326..340 /note= "T-cell epitope peptide"
FT peptide 335..346 /note= "T-cell epitope peptide"
PN J07118295-A.
PD C9-MAY-1995.
PF 20-OCT-1993: 262626.
PR 20-OCT-1993: JP-262626.
PA (MEIP) MEIJI MILK PROD CO LTD.
DR WPI: 95-203834/27.
PI New cryptomeria pollen allergen T-cell epitope peptide - used for
PI prevention, treatment and investigation of Japanese cedar pollinosis.
PS Disclosure: Figs 1-2: 8pp; Japanese.
CC R75388 is the Japanese cedar pollen allergen Cryj I, from which the
CC T-cell epitope peptides R892, R89295 were derived. The peptides
CC can be used for the prevention and treatment of cryptomeria
CC pollinosis, and also for the investigation of pollinosis.
SQ Sequence 353 AA:
Query Match 100.0%; Score 94; DB 1; Length 353;
Best Local Similarity 100.0%; Pred. No. 2.88e-03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
DB 108 VFIKRVSNVILHG 120
QY 1 VFIKRVSNVILHG 13
RESULT 9
ID R81587 standard; Protein; 353 AA.
AC R81587 (first entry)
DE Cedar pollen allergen B.
KW Cedar; pollen allergen; immunoglobulin E; IgE; T-cell epitope;
KW antibody; pollinosis; therapy; immunotherapy.
OS Cryptomeria japonica.
PN EP-700929-A2.
PC 13-MAR-1996.
PF 08-SEP-1995: 306295.
PR 10-SEP-1994: JP-242137.
PR 14-JUL-1995: JP-200221.
PR 14-JUL-1995: JP-200204.
PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.
PI Hino K, Saito S, Taniguchi Y;
DR WPI: 96-140976/15.
PT New peptide(s) derived from cedar pollen allergens - activate
PT allergen-specific T-cells, but not allergen-specific IgE antibodies.
PI used for treating cedar pollinosis
PS Claim 5: Page 31-32: 36pp; English.
CC Synthetic peptides based on portions of cedar pollen allergens A
CC (R81586) and B (R81587) were tested for their ability to activate
CC cedar allergen-specific T-cells, but not allergen-specific IgE
CC antibodies. 6 peptides (R81580-R8185) were identified as T-cell
CC epitopes. These peptides, plus subsequences (R81573-79) essential
CC for T-cell recognition, and homologous peptides (R81588-96) can
CC be used as immunotherapeutic agents to treat or prevent cedar
CC pollinosis, avoiding side-effects such as anaphylaxis.

Sequence 353 AA:
 Query Match 100.0% Score 94: DB 1: Length 353:
 Best Local Similarity 100.0% Pred. No. 2.88e-03:
 Matches 13: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

DB 108 VFIRKRVSNV:HG 120
 ||||| |||||
 QY 1 VFIRKRVSNV:HG 13

RESULT 10
 ID R31937 standard: Protein: 374 AA.
 AC R31937:
 DT 03-JUN-1993 (first entry)
 DE Cry j 1: Japanese cedar pollen; allergen; antigen; allergy: B cell; T cell.
 KW Cryptomeria japonica.
 FH Key Location/Qualifiers
 FI Peptide 1..21 /note= "signal peptide"
 FT protein 22..374 /note= "mature Cry j 1"
 ST WO9300213-A.
 SN 21-JAN-1993:
 SC 10-JUL-1992: U05661.
 PR 12-JUL-1991: US-729134.
 PR 15-JUL-1991: US-730452.
 PA (IMMUNO) IMMUNOLOGIC PHARM CORP.
 PI Bond JF, Griffith IC, Pollock J.
 DR WP1: 93-045434/05.
 OR N-PSDB: Q35304.
 PT Nucleic acid sequence encoding Cryptomeria japonica allergen -
 for the diagnosis treatment and prevention of allergic reactions
 to Japanese cedar pollen
 PS Claim 1: Page 42: 69pp: English.
 CC Fresh pollen and staminate cone samples were collected from a single
 Cryptomeria japonica (Japanese cedar) tree. RNA was prep'd. and used
 to synthesize cDNA. The cDNA was subjected to successive rounds of
 PCR to yield a full length Cry j 1 clone. Cry j 1 or an antigenic
 fragment of it may be used for detecting, treating and preventing an
 allergic response to Japanese cedar pollen allergen. It is capable of
 modifying both the B and T cell response to Cry j 1 and T cell response
 to a Cry j 1 antigen.
 SC Sequence 374 AA:

Query Match 100.0% Score 94: DB 1: Length 374:
 Best Local Similarity 100.0% Pred. No. 2.88e-03:
 Matches 13: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

DB 129 VFIRKRVSNV:HG 141
 ||||| |||||
 QY 1 VFIRKRVSNV:HG 13

RESULT 11
 ID R82490 standard: Protein: 374 AA.
 AC R82490:
 DT 15-APR-1996 (first entry)
 DE Cry j 1: Japanese Cedar pollen allergen.
 KW Cry j 1: Japanese cedar pollen allergen; modified: drug production;
 allergen; Cryptomeria japonica.
 FH Key Location/Qualifiers
 FI Peptide 1..21 /note= "signal peptide"
 FT protein 22..374 /note= "mature Cry j 1"
 ST WO9527786-A.
 SN 19-OCT-1995:
 PR 06-APR-1995: U04249.
 PR 08-APR-1994: US-226248.
 PR 06-DEC-1994: US-350325.
 PA (IMMUNO) IMMUNOLOGIC PHARM CORP.
 PI Chen X, Evans S, Franzen HM, Kuo M, Powers SP,
 PI Skaked Z.
 DR WP1: 95-366391/47.
 OR N-PSDB: T04249.

PT Modified Cryptomeria japonica (Cry j 1) peptide(s) - useful for
 treating allergy to Japanese cedar pollen allergen or
 immunologically cross reactive allergens
 PS Disclosure: Figure 1: 60pp: English.
 CC Novel peptides of cry j 1 have been modified as a part of a
 preformulation scheme to develop an optimised drug product for
 therapeutic treatment of human suffering from allergy to Japanese
 cedar pollen allergen or an allergen which is immunologically cross
 reactive with Japanese cedar pollen allergen. Such modified peptides
 possess certain characteristics which render them particularly
 suitable for drug product formulation. Peptide fragments of Cry j 1
 modified and unmodified, are given in R82491-R82525.
 SC Sequence 374 AA:

Query Match 100.0% Score 94: DB 1: Length 374:
 Best Local Similarity 100.0% Pred. No. 2.88e-03:
 Matches 13: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

DB 129 VFIRKRVSNV:HG 141
 ||||| |||||
 QY 1 VFIRKRVSNV:HG 13

RESULT 12
 ID R45541 standard: Protein: 374 AA.
 AC R45541:
 DT 13-JUL-1994 (first entry)
 DE Cry j 1 pollen allergen
 KW Japanese cedar; detection: allergy; treatment: diagnosis;
 T cell epitope; sensitivity.
 CS Cryptomeria japonica.
 FH Key Location/Qualifiers
 FI Peptide 1..21 /note= "signal peptide"
 FT peptide 22..374 /note= "mature peptide"
 ST WO9401560-A.
 SN 20-JAN-1994:
 PR 15-JAN-1993: U00139.
 PR 10-JUL-1992: WO-U05661.
 PR 01-SEP-1992: US-938990.
 PA (IMMUNO) IMMUNOLOGIC PHARM CORP.
 PI Bond JF, Garman RD, Griffith IC, Pollock J.
 DR WP1: 94-035066/04.
 OR N-PSDB: Q55271.
 PT Antigens derived from Japanese cedar pollen allergen Cry j 1
 contain at least two T cell epitope(s). used to treat or diagnose
 allergy
 PS Disclosure: Fig 4: 137pp: English
 CC The sequence is that of the Japanese cedar pollen allergen
 Cry j 1 which contains at least two T cell epitopes. Peptide
 antigens derived from it can be used for the treatment and
 diagnosis of allergies associated with Japanese cedar pollen.
 CC The peptides have enhanced therapeutic properties but reduced
 side effects compared to naturally occurring allergens.
 SC Sequence 374 AA:

Query Match 100.0% Score 94: DB 1: Length 374:
 Best Local Similarity 100.0% Pred. No. 2.88e-03:
 Matches 13: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

DB 129 VFIRKRVSNV:HG 141
 ||||| |||||
 QY 1 VFIRKRVSNV:HG 13

RESULT 13
 ID R50166 standard: Protein: 374 AA.
 AC R50166:
 DT 24-MAR-1995 (first entry)
 DE Japanese cedar pollen antigen Cryj1
 KW Japanese cedar; pollen antigen; allergen; Cryj1; sugi; pollinosis.
 CS Cryptomeria japonica.

DE Key Location/Qualifiers
 DT peptide /label: signal_peptide
 KW protin 22, 374
 OS /label: mature_CryJ1
 PN J06197764.A.
 PD 19-JUL-1994
 PF 07-JAN-1993: C00139.
 PR 07-JAN-1993: JP-001116.
 PA (METS) METS: SEIKA KAISHA.
 LR WP1: 94-26866/13
 RP WP-ESCB: 072601.
 PT Sugi (Japanese cedar) pollen antigen CryJ1 is useful for
 PS diagnosis, treatment and prevention of sugi pollinosis.
 PS Claim 2: Pate 5-7: pp: Japanese.
 CC The coding sequence for the Japanese cedar ("sugi") pollen allergen
 CC CryJ1 was isolated from a cDNA library prepared from polyA mRNA. All
 CC or part of the CryJ1 protein can be used for diagnosis, treatment
 CC and prevention of sugi pollinosis.
 SQ Sequence 171 AA:

Query Match: 100.0% Score 941 DB 1: Length 374:
 Best Local Similarity 100.0% Pred. No. 2 896-03:
 Matches 13: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

4 129 VKKRVSNVTHG 14:
 11111111
 5 1 VKKRVSNVTHG 13

RESULT 14
 IC W57750 standard: peptide: 15 AA.

AC W57750
 DT 17-SEP-1998 (first entry)
 KW Residues 165-220 of Cry J1.
 SW Cry J1, Japanese Cedar pollen antigen; allergy: immunotherapy;
 KW HLA class II molecule.
 OS Cryptomeria japonica.
 PN W098209.2.A.
 PD 22-MAY-1998
 PF 12-N-V-1997: C04129.
 PR 11-N-V-1996: JP-402053.
 PA (METS) METS: HLK PECO CO LTD.
 LR Gairiki K, Kuno K, Kume A, Soga T:
 WP1: 98-29761/25.
 PT Peptides derived from Japanese cedar pollen antigens are
 PS immunotherapeutic agents - useful for allergy treatment and typing
 PS HLA class II molecules in allergy sufferers
 PS Claim 1: Par 26: 5-11: Japanese
 CC This sequence represents residues 165-220 of the Cry J1 protein and
 CC is a peptide of the invention. The peptides are derived from Japanese
 CC cedar pollen antigens and are used as immunotherapeutic agents in the
 CC treatment of allergy. The peptides can be used for identification and
 CC typing of the particular HLA class II molecules in an allergy sufferer.
 CC and also for peptide immunotherapy of an allergy. Using these peptides
 CC the immunotherapy can be targeted more specifically to the requirements
 CC of the individual patient, allowing more effective treatment of an
 CC allergy, including those patients for whom treatment with a conventional
 CC immunotherapeutic agent is ineffective.
 SQ Sequence 15 AA:

Query Match: 87.2% Score 82: DB 1: Length 15:
 Best Local Similarity 92.3% Pred. No. 7 228-02:
 Matches 12: Conservative 0: Mismatches 1: Indels 0: Gaps 0:

DE 3 VKKRVSNVTHG 15
 111111111
 QY 1 VKKRVSNVTHG 13

RESULT 15
 IC R45551 standard: Protein: 20 AA.
 AC R45551

DT 13-JUL-1994 (first entry)
 DE Cry J1 pollen allergen peptide C01-12.
 KW Japanese cedar; detection; allergy; treatment; diagnosis;
 KW I cell epitope; sensitivity.
 OS Cryptomeria japonica.
 PN W09401560.A.
 PD 20-JAN-1994: C00139.
 PF 15-JAN-1993: WO-005661.
 PR 10-JUL-1992: WO-005661.
 PA (IMMUNO) IMMUNOLOGIC PHARM CORP
 RI Bond JF, Garman RD, Griffith LR, Kuo M, Pollock J:
 WP1: 94-035086/04.
 PT Antigens derived from Japanese cedar pollen allergen Cry J1
 PT contain at least two T cell epitope(s), used to treat or prevent
 PT allergy
 PS Claim 1: Fig 13: 137pp: English.
 CC The sequence is that of an isolated peptide of the Japanese cedar
 CC pollen allergen Cry J1 (amino acids 111-133). The peptide C01-12
 CC can be used for the treatment and diagnosis of allergies associated
 CC with Japanese cedar pollen. It has enhanced therapeutic properties
 CC but reduced side effects for use to naturally occurring allergies.
 SQ Sequence 20 AA:

Query Match: 73.4% Score 59: DB 1: Length 20:
 Best Local Similarity 100.0% Pred. No. 2 124-01:
 Matches 10: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

DE 1 KRVSNVTHG 10
 1111111111
 QY 4 KRVSNVTHG 13

Search completed: Mon Jun 19 11:11:40 2000
 Job time: 6 secs.

.....
[W][O][R][D] [I][M]
.....

Release J.L.A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Tue Jun 20 13:37:24 2000: MasPar time 2.28 Seconds
Tabular output not generated. 82.449 Million cell updates/sec

Title: >US-09-142-524A-13
Description: (1-13) from US09142524A.pep
Perfect Score: 94
Sequence: 1 VFIKRVSNVTHG 13

Scoring table: PAM 150
Gap 15

Searched: 145341 seqs, 14437480 residues

Post-processing: Minimum Match 0%
Listing first: 45 summaries

Database: a-issued
1:5A_COMB 2:5B_COMB 3:6_COMB 4:PCI_COMB 5:backfiles1

Statistics: Mean 17.285; Variance 56.805; scale 0.304

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the results being printed, and is derived by analysis of the total score distribution.

| SUMMARIES | | | |
|------------|-------|--------------------------|---|
| Result No. | Score | Query Match length DB ID | Description Pred. No. |
| 1 | 56.5 | 255 | US-08-459- Sequence 4, Applicatio 3.56e-01 |
| 2 | 56.5 | 255 | US-08-459- Sequence 4, Applicatio 3.56e-01 |
| 3 | 56.5 | 428 | US-09-120- Sequence 2, Applicatio 3.56e-01 |
| 4 | 56.5 | 428 | US-08-326- Sequence 2, Applicatio 3.56e-01 |
| 5 | 52.3 | 132 | PCI-US95-1 Sequence 80, Applicati 9.20e-01 |
| 6 | 52.3 | 132 | US-08-086- Sequence 80, Applicati 9.20e-01 |
| 7 | 52.3 | 132 | US-08-290- Sequence 80, Applicati 9.20e-01 |
| 8 | 52.3 | 132 | US-08-458- Sequence 14, Applicati 9.20e-01 |
| 9 | 51.5 | 295 | US-08-504- Sequence 12, Applicati 1.16e-02 |
| 10 | 51.5 | 297 | US-08-504- Sequence 12, Applicati 1.16e-02 |
| 11 | 51.5 | 625 | US-08-365- Sequence 13, Applicati 1.16e-02 |
| 12 | 50.3 | 316 | US-08-846- Sequence 13, Applicatio 1.45e-02 |
| 13 | 50.3 | 317 | US-08-951- Sequence 2, Applicatio 1.45e-02 |
| 14 | 50.3 | 317 | US-09-056- Sequence 2, Applicatio 1.45e-02 |
| 15 | 50.3 | 317 | US-08-518- Sequence 2, Applicatio 1.45e-02 |
| 16 | 50.3 | 908 | US-08-249- Sequence 2, Applicatio 1.45e-02 |
| 17 | 49.5 | 299 | US-08-504- Sequence 10, Applicati 1.81e-02 |
| 18 | 48.5 | 85 | US-08-659- Sequence 3, Applicatio 2.27e-02 |
| 19 | 48.5 | 85 | US-08-480- Sequence 3, Applicatio 2.27e-02 |
| 20 | 48.5 | 111 | US-08-111- Sequence 23, Applicati 2.27e-02 |
| 21 | 48.5 | 160 | US-08-162- Sequence 14, Applicati 2.27e-02 |
| 22 | 48.5 | 218 | US-08-162- Sequence 4, Applicatio 2.27e-02 |
| 23 | 48.5 | 218 | US-07-607- Sequence 4, Applicatio 2.27e-02 |

STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE: Bacillus cereus
 ORGANISM: Bacillus cereus
 IMMEDIATE SOURCE: B-LACTAMASE
 CLONE: B-LACTAMASE
 SEQUENCE 255 AA: 27847 MW: 349604 CN:

Query Match 59.6% Score 56: DB 1: Length 255:
 Best Local Similarity 70.0%: Pred. No. 3,66e-01:
 Matches 7: Conservative 2: Mismatches 1: Indels 0: Gaps 0:

DB 103 FOKRVIDVII 112
 I I I I I I I I
 QY 2 FIKRVSNVII 11

RESULT 2
 ID US-09-459-263-4 STANDARD: PRT: 255 AA.
 XX
 AC xxxxxx
 XX
 DT
 XX

Sequence 4: Application US/04459263

Sequence 4: Application US/04459263
 Patent No. 5792642
 GENERAL INFORMATION:
 APPLICANT: RASMUSSEN, Beth A
 APPLICANT: TALSY, Francis P
 APPLICANT: GLJZMAN, Yakov
 TITLE OF INVENTION: DNA SEQUENCES AND AMINO ACID SEQUENCES
 TITLE OF INVENTION: OF CLASS B BETA-LACTAMASE ENZYMES FROM BACTERIOIDES
 TITLE OF INVENTION: FRAGLIS
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Darby & Darby PC
 STREET: 805 Third Avenue
 CITY: New York
 STATE: NY
 COUNTRY: US
 ZIP: 10022

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/8,455,273
 FILING DATE: 02-JUN-1995
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Robinson, Joseph R
 REGISTRATION NUMBER: 33,448
 REFERENCE/DOCKET NUMBER: C646/1B026-052
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-527-7700
 TELEFAX: 212-753-6237
 TELEX: 236687

INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 255 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 ORIGINAL SOURCE: Bacillus cereus
 ORGANISM: Bacillus cereus
 IMMEDIATE SOURCE: B-LACTAMASE
 CLONE: B-LACTAMASE
 SEQUENCE 255 AA: 27847 MW: 349604 CN:

Query Match 59.6% Score 56: DB 1: Length 255:
 Best Local Similarity 70.0%: Pred. No. 3,66e-01:
 Matches 7: Conservative 2: Mismatches 1: Indels 0: Gaps 0:

DB 103 FOKRVIDVII 112
 I I I I I I I I
 QY 2 FIKRVSNVII 11

RESULT 3
 ID US-09-120-353-2 STANDARD: PRT: 408 AA.
 XX
 AC xxxxxx
 XX
 DT
 XX

Sequence 2: Application US/09120054

Sequence 2: Application US/09120054
 Patent No. 5932454
 GENERAL INFORMATION:
 APPLICANT: Baltz, Richard H.
 APPLICANT: Solenberg, Patricia J.
 APPLICANT: Treadway, Patti J.
 TITLE OF INVENTION: Glucosyltransferase Gene Jute From
 TITLE OF INVENTION: Amycolatopsis orientalis
 NUMBER OF SEQUENCES: 3
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Eli Lilly and Company
 STREET: Lilly Corporate Center
 CITY: Indianapolis
 STATE: Indiana
 COUNTRY: USA
 ZIP: 46285
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/9,120,054
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/945,258
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Webster, Thomas L.
 REGISTRATION NUMBER: 44,571
 REFERENCE/DOCKET NUMBER: P01474
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 317-276-3334
 TELEFAX: 317-276-3861
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 408 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE 408 AA: 43042 MW: 779745 CN:

Query Match 59.6% Score 56: DB 2: Length 408:
 Best Local Similarity 53.8%: Pred. No. 3,66e-01:
 Matches 7: Conservative 4: Mismatches 2: Indels 0: Gaps 0:

DB 298 VLFRRVAAVIHG 310
 I I I I I I I I
 QY 1 VFIKRVSNVII 13

RESULT 4
 ID US-08-926-258-2 STANDARD: PRT: 408 AA.

CC ADDRESSEE: MORGAN & FINNEGAN
 CC STREET: 345 PARK AVENUE
 CC CITY: NEW YORK
 CC STATE: NEW YORK
 CC COUNTRY: USA
 CC ZIP: 10154
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: FLOPPY DISK
 CC COMPUTER: IBM PC COMPATIBLE
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: WORDPERFECT 5.1
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/086.428B
 CC FILING DATE: 29-JUN-1993
 CC CLASSIFICATION: 435
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: RICHARD W. BORK
 CC REGISTRATION NUMBER: 36.459
 CC REFERENCE/DOCKET NUMBER: 2026-4270
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (212) 758-4800
 CC TELEFAX: (212) 751-6849
 CC TELEX: 421792
 CC INFORMATION FOR SEQ ID NO: 80:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 192 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS: unknown
 CC TOPOLOGY: unknown
 CC ORIGINAL SOURCE:
 CC ORGANISM: homosapiens
 CC INDIVIDUAL ISOLATE: US10
 CC SEQUENCE: 192 AA; 20855 MW; 200760 CN;

Query Match: 55.3% Score 52: 192 Length 192:
 Best Local Similarity 50.0% Pred. No. 9: 20e-01:
 Matches 5: Conservative 5: Mismatches 0: Indels 0: Gaps 0:

DB 145 YVNRVPEVIL 154
 QY 2 FIKRVSNVII 11

PERJUL 7
 CC US-08 290-665A-60 STANDARD: PRT: 192 AA.
 CC xxxxxx

DE Sequence 80, Application US/08/086.428B
 XX Sequence 80, Application US/08/086.428B
 CC Patent No. 5882552
 CC GENERAL INFORMATION:
 CC APPLICANT: BUKH, J., MILLER, R. H. AND
 CC TITLE OF INVENTION: NUCLEOTIDE AND DERIVED
 CC TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
 CC TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
 CC TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
 CC TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
 CC NUMBER OF SEQUENCES: 263
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: MORGAN & FINNEGAN
 CC STREET: 345 PARK AVENUE
 CC CITY: NEW YORK
 CC STATE: NEW YORK
 CC COUNTRY: USA
 CC ZIP: 10154
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: FLOPPY DISK
 CC COMPUTER: IBM PC COMPATIBLE

CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: WORDPERFECT 5.1
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: 08/290.665A
 CC FILING DATE: 15-AUG-1
 CC CLASSIFICATION: 435
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: RICHARD W. BORK
 CC REGISTRATION NUMBER: 36.459
 CC REFERENCE/DOCKET NUMBER: 2026-4116
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (212) 758-4800
 CC TELEFAX: (212) 751-6849
 CC TELEX: 421792
 CC INFORMATION FOR SEQ ID NO: 80:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 192 amino acids
 CC TYPE: amino acid
 CC STRANDEDNESS: unknown
 CC TOPOLOGY: unknown
 CC ORIGINAL SOURCE:
 CC ORGANISM: homosapiens
 CC INDIVIDUAL ISOLATE: US10
 CC SEQUENCE: 192 AA; 20855 MW; 200760 CN;

Query Match: 55.3% Score 52: 192 Length 192:
 Best Local Similarity 50.0% Pred. No. 9: 20e-01:
 Matches 5: Conservative 5: Mismatches 0: Indels 0: Gaps 0:

DB 145 YVNRVPEVIL 154
 QY 2 FIKRVSNVII 11

RESULT 2
 ID US-08-468-570-80 STANDARD: PRT: 192 AA.
 XX xxxxxx

DE Sequence 80, Application US/08/086.428B
 XX Sequence 80, Application US/ 8458570
 CC Patent No. 5873562
 CC GENERAL INFORMATION:
 CC APPLICANT: BUKH, J., MILLER, R. H. AND
 CC TITLE OF INVENTION: NUCLEOTIDE AND DERIVED
 CC TITLE OF INVENTION: AMINO ACID SEQUENCES OF THE ENVELOPE 1 AND
 CC TITLE OF INVENTION: CORE GENES OF ISOLATES OF HEPATITIS C VIRUS
 CC TITLE OF INVENTION: AND THE USE OF REAGENTS DERIVED FROM THESE
 CC TITLE OF INVENTION: SEQUENCES IN DIAGNOSTIC METHODS AND VACCINES
 CC NUMBER OF SEQUENCES: 159
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: MORGAN & FINNEGAN
 CC STREET: 345 PARK AVENUE
 CC CITY: NEW YORK
 CC STATE: NEW YORK
 CC COUNTRY: USA
 CC ZIP: 10154
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: FLOPPY DISK
 CC COMPUTER: IBM PC COMPATIBLE
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: WORDPERFECT 5.1
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: 08/468.570
 CC FILING DATE: 6-JUN-93
 CC CLASSIFICATION: 424
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/086.428
 CC FILING DATE: 29-JUN-1993

CC CLASSIFICATION: 424
CC ATTORNEY/AGENT INFORMATION:
CC NAME: RICHARD W. BORK
CC REGISTRATION NUMBER: 36,459
CC REFERENCE/DOCKET NUMBER: 2026-4070US1
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (212) 758-4800
CC TELEFAX: (212) 751-6849
CC TELEX: 421792
CC INFORMATION FOR SEQ ID NO: 80:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 192 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: unknown
CC TOPOLOGY: unknown
CC ORIGIN: SOURCE:
CC ORGANISM: homosapiens
CC INDIVIDUAL ISOLATE: US10
CC SEQUENCE: 192 AA; 20855 MW; 200760 CN;

Query Match: 55.3%; Score 52; DB 2; Length 192;
Best Local Similarity 50.0%; Pred. NO. 9.20e+01;
Matches 5; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

DB 145 YVMRPEVII 154
QY 2 FIKRVSNVII 11

RESULT 5
ID US-08-504-459-14 STANDARD: PRT: 295 AA.

XX

AC xxxxxx

XX

DT

XX

Sequence 14, Application US/08504459

Sequence 14, Application US/08504459

Patent No. 5922563

GENERAL INFORMATION:

APPLICANT: Alderete, John F.

TITLE OF INVENTION: Adhesin Genes and Proteins Involved in

TITLE OF INVENTION: Trichomonas Vaginalis Cyoadherence

NUMBER OF SEQUENCES: 30

CORRESPONDENCE ADDRESS:

ADDRESSEE: Arnold, White & Durkee

STREET: P.O. Box 4433

CITY: Houston

STATE: Texas

COUNTRY: US

ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/504,459

FILING DATE: Concurrently herewith

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Wilson, Mark B.

REGISTRATION NUMBER: 37,259

REFERENCE/DOCKET NUMBER: US/08/504,459

TELECOMMUNICATION INFORMATION:

TELEPHONE: 512/418-3000

TELEFAX: 512/474-7577

TELEX: N/A

INFORMATION FOR SEQ ID NO: 14:

SEQUENCE CHARACTERISTICS:

LENGTH: 295 amino acids

TYPE: amino acid

CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
SQ SEQUENCE 295 AA; 30887 MW; 423730 CN;

Query Match: 54.3%; Score 51; DB 2; Length 295;
Best Local Similarity 38.5%; Pred. NO. 1.16e+02;
Matches 5; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

DB 1 LFIDKDTKVVIGG 13
QY 1 VFIRKRVSNVII 13

RESULT 10
ID US-08-504-459-12 STANDARD: PRT: 297 AA.
XX
AC xxxxxx
XX
DT
XX
XX
DE Sequence 12, Application US/08504459
XX
CC Sequence 12, Application US/08504459
CC Patent No. 5922563
CC GENERAL INFORMATION:
CC APPLICANT: Alderete, John F.
CC TITLE OF INVENTION: Adhesin Genes and Proteins Involved in
CC TITLE OF INVENTION: Trichomonas Vaginalis Cyoadherence
CC NUMBER OF SEQUENCES: 30
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Arnold, White & Durkee
CC STREET: P.O. Box 4433
CC CITY: Houston
CC STATE: Texas
CC COUNTRY: US
CC ZIP: 77210
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/504,459
CC FILING DATE: Concurrently herewith
CC CLASSIFICATION: 435
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Wilson, Mark B.
CC REGISTRATION NUMBER: 37,259
CC REFERENCE/DOCKET NUMBER: US/08/504,459
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 512/418-3000
CC TELEFAX: 512/474-7577
CC TELEX: N/A
CC INFORMATION FOR SEQ ID NO: 12:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 297 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
SQ SEQUENCE 297 AA; 31082 MW; 429472 CN;

Query Match: 54.3%; Score 51; DB 2; Length 297;
Best Local Similarity 38.5%; Pred. NO. 1.16e+02;
Matches 5; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

DB 3 LFIDKDTKVVIGG 15

QY 1 VFIRKRVSNVII 13

RESULT 11

ID US-08-365-981-13 STANDARD: PRT: 625 AA.

XX

```

AC      xxxxxx
AD
AE
AF
AG
AH
AI
AJ
AK
AL
AM
AN
AO
AP
AQ
AR
AS
AT
AU
AV
AW
AX
AY
AZ
BA
BB
BC
BD
BE
BF
BG
BH
BI
BJ
BK
BL
BM
BN
BO
BP
BQ
BR
BS
BT
BU
BV
BW
BX
BY
BZ
CA
CB
CC
CD
CE
CF
CG
CH
CI
CJ
CK
CL
CM
CN
CO
CP
CQ
CR
CS
CT
CU
CV
CW
CX
CY
CZ
DA
DB
DC
DD
DE
DF
DG
DH
DI
DJ
DK
DL
DM
DN
DO
DP
DQ
DR
DS
DT
DU
DV
DW
DX
DY
EZ
FA
FB
FC
FD
FE
FF
FG
FH
FI
FJ
FK
FL
FM
FN
FO
FP
FQ
FR
FS
FT
FU
FV
FW
FX
FY
FZ
GA
GB
GC
GD
GE
GF
GG
GH
GI
GJ
GK
GL
GM
GN
GO
GP
GQ
GR
GS
GT
GU
GV
GW
GX
GY
GZ
HA
HB
HC
HD
HE
HF
HG
HH
HI
HJ
HK
HL
HM
HN
HO
HP
HQ
HR
HS
HT
HU
HV
HW
HX
HY
HZ
IA
IB
IC
ID
IE
IF
IG
IH
II
IJ
IK
IL
IM
IN
IO
IP
IQ
IR
IS
IT
IU
IV
IW
IX
IY
IZ
JA
JB
JC
JD
JE
JF
JG
JH
JI
JJ
JK
JL
JM
JN
JO
JP
JQ
JR
JS
JT
JU
JV
JW
JX
JY
JZ
KA
KB
KC
KD
KE
KF
KG
KH
KI
KJ
KK
KL
KM
KN
KO
KP
KQ
KR
KS
KT
KU
KV
KW
KX
KY
KZ
LA
LB
LC
LD
LE
LF
LG
LH
LI
LJ
LK
LL
LM
LN
LO
LP
LQ
LR
LS
LT
LU
LV
LW
LX
LY
LZ
MA
MB
MC
MD
ME
MF
MG
MH
MI
MJ
MK
ML
MM
MN
MO
MP
MQ
MR
MS
MT
MU
MV
MW
MX
MY
MZ
NA
NB
NC
ND
NE
NF
NG
NH
NI
NJ
NK
NL
NM
NN
NO
NP
NQ
NR
NS
NT
NU
NV
NW
NX
NY
NZ
OA
OB
OC
OD
OE
OF
OG
OH
OI
OJ
OK
OL
OM
ON
OO
OP
OQ
OR
OS
OT
OU
OV
OW
OX
OY
OZ
PA
PB
PC
PD
PE
PF
PG
PH
PI
PJ
PK
PL
PM
PN
PO
PP
PQ
PR
PS
PT
PU
PV
PW
PX
PY
PZ
QA
QB
QC
QD
QE
QF
QG
QH
QI
QJ
QK
QL
QM
QN
QO
QP
QQ
QR
QS
QT
QU
QV
QW
QX
QY
QZ
RA
RB
RC
RD
RE
RF
RG
RH
RI
RJ
RK
RL
RM
RN
RO
RP
RQ
RR
RS
RT
RU
RV
RW
RX
RY
RZ
SA
SB
SC
SD
SE
SF
SG
SH
SI
SJ
SK
SL
SM
SN
SO
SP
SQ
SR
SS
ST
SU
SV
SW
SX
SY
SZ
TA
TB
TC
TD
TE
TF
TG
TH
TI
TJ
TK
TL
TM
TN
TO
TP
TQ
TR
TS
TT
TU
TV
TW
TX
TY
TZ
UA
UB
UC
UD
UE
UF
UG
UH
UI
UJ
UK
UL
UM
UN
UO
UP
UQ
UR
US
UT
UU
UV
UW
UX
UY
UZ
VA
VB
VC
VD
VE
VF
VG
VH
VI
VJ
VK
VL
VM
VN
VO
VP
VQ
VR
VS
VT
VU
VV
VW
VX
VY
VZ
WA
WB
WC
WD
WE
WF
WG
WH
WI
WJ
WK
WL
WM
WN
WO
WP
WQ
WR
WS
WT
WU
WV
WW
WX
WY
WZ
XA
XB
XC
XD
XE
XF
XG
XH
XI
XJ
XK
XL
XM
XN
XO
XP
XQ
XR
XS
XT
XU
XV
XW
XX
XY
XZ
YA
YB
YC
YD
YE
YF
YG
YH
YI
YJ
YK
YL
YM
YN
YO
YP
YQ
YR
YS
YT
YU
YV
YW
YX
YY
YZ
ZA
ZB
ZC
ZD
ZE
ZF
ZG
ZH
ZI
ZJ
ZK
ZL
ZM
ZN
ZO
ZP
ZQ
ZR
ZS
ZT
ZU
ZV
ZW
ZX
ZY
ZZ

```

Sequence 13, Application US/08365981

Sequence 13, Application US/08365981

Patent No. 5583330

GENERAL INFORMATION:

APPLICANT: ROBERT DICKSON et al.

TITLE OF INVENTION: METHOD FOR OBTAINING

TITLE OF INVENTION: ANTI-FUNGAL AND ANTIBACTERIAL COMPOUNDS THAT TARGET THE

TITLE OF INVENTION: FIRST COMMITTED STEP IN SPHINGOLIPID LONG-CHAIN

TITLE OF INVENTION: BASE BIOSYNTHESIS

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS: 15

ADDRESSEE: LOWE, PRICE, IEBLANC & PECKER

STREET: STE. 300, 99 CANAL CENTER PLAZA

CITY: Alexandria

STATE: Virginia

COUNTRY: USA

ZIP: 22314

COMPUTER READABLE FORM:

MEDIA TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: DOS Text File

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/365,981

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/906,899

FILING DATE: 06/30/92

ATTORNEY/AGENT INFORMATION:

NAME: D.J. Mills

REGISTRATION NUMBER: 34506

REFERENCE/DOCKET NUMBER: 434-029

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 684-1111

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 626

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: polypeptide

SEQUENCE: 625 AA: 69612 MW: 2352240 CN.

Query Match 54.1% Score 511.14 11 Length 626

Best Local Similarity 41.7% Pctd. No. 1160-041

Matches 5: Conservative 4: Mismatches 1: Indels 0: Gaps 0:

DB 250 FIKRSGDVVXA 261

QY 1: 111111

QY 2 FIKRVSNNVTHG 13

Sequence 9, Application US/08846762A

Sequence 9, Application US/08846762A

Patent No. 5994072

GENERAL INFORMATION:

APPLICANT: Lam, Joseph S.

APPLICANT: Burrows, Lori

APPLICANT: Chatter, Deborah

APPLICANT: de Kievit, Teresa

TITLE OF INVENTION: No. 5994072el Proteins Involved in the Synthesis and Assem

TITLE OF INVENTION: of O-Antigen in Pseudomonas Aeruginosa

FILE REFERENCE: 6580-089

CURRENT APPLICATION NUMBER: US/08/846,762A

CURRENT FILING DATE: 1997-04-30

NUMBER OF SEQ ID NOS: 100

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 9

LENGTH: 316

TYPE: PRT

ORGANISM: Pseudomonas aeruginosa

SEQUENCE 316 AA: 33805 MW: 521677 CN:

Query Match 53.2% Score 501.08 21 Length 316

Best Local Similarity 50.0% Pctd. No. 11450-021

Matches 5: Conservative 4: Mismatches 1: Indels 0: Gaps 0:

DB 5 FIKRVARVLDV 14

QY 11111111

QY 2 FIKRVSNNVTHG 11

Sequence 2, Application US/08951889

Sequence 2, Application US/08951889

Patent No. 6008032

GENERAL INFORMATION:

APPLICANT: Mathur, E. et al.

TITLE OF INVENTION: Carboxymethyl Cellulase from

TITLE OF INVENTION: Thermotoga Maritima

NUMBER OF SEQUENCES: 4

CORRESPONDENCE ADDRESS:

ADDRESSEE: CARELLA, BYRGE, SA'N, GILFILLAN,

STREET: 6 BECKER FARM ROAD

CITY: ROSELAND

STATE: NEW JERSEY

COUNTRY: USA

ZIP: 07068

COMPUTER READABLE FORM:

MEDIA TYPE: 3.5 INCH DISKETTE

COMPUTER: IBM PS/2

OPERATING SYSTEM: MS-DOS

SOFTWARE: WORD PERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/951,889

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/518,615

FILING DATE: August 23, 1995

ATTORNEY/AGENT INFORMATION:

NAME: FERRARO, GREGORY D.

REGISTRATION NUMBER: 36,134

REFERENCE/DOCKET NUMBER: 331400-20

TEL: COMMUNICATION INFORMATION:

TELEPHONE: 201-994-1701

TELEFAX: 201-994-1744

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 317 AMINO ACIDS

TYPE: AMINO ACID

STRANDEDNESS:

TOPOLOGY: LINEAR

MOLECULE TYPE: PROTEIN

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

| protein - protein database search, using Smith-Waterman algorithm | Run on: | Mon Jun '9 16:28:19 2000: | Maxpar time 14.74 Seconds |
|---|---------|---------------------------|---------------------------------|
| | | | 88.845 Million cell updates/sec |

```
>JS-09-142-524A-13
Description: (1-13) from US09142524A pep
Perfect Score: 94
Sequence: 1 VFIKRVSNVIHG 13
```

Scoring table:

| |
|---------|
| PAM 150 |
| Gap 15 |

Searched: 721208 seqs, 100765575 residues

Post-processing: Minimum Match 08
Listing first 45 summaries

Database: a-pending

| | | | | | | | | |
|-----------|--------|--------|--------|---------|-----------|----------|--------|--------|
| 1:PCI | 2:06 | 3:060 | 4:07 | 5:08C | 6:081 | 7:082 | 8:083 | 9:084A |
| 10:084B | 11:085 | 12:086 | 13:087 | 14:088 | 15:089 | 16:090 | 17:091 | |
| 18:092 | 19:093 | 20:094 | 21:095 | 22:NEWP | 23:NEWU60 | 24:NEWJ8 | | |
| 25:NEWK19 | | | | | | | | |

Statistics: N=30; Variance 54.778; scale 0.375

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query | | Len _q | DB | ID | Description | Pred No. |
|------------|-------|-------|-----|------------------|------------|--------------|-------------|----------|
| | | Match | | | | | | |
| 1 | 94 | 100.0 | 20 | 8 | US-08-330- | Sequence 36, | Applicat | 3.61e-03 |
| 2 | 94 | 100.0 | 20 | 8 | US-08-236- | Sequence 36, | Applicat | 3.61e-03 |
| 3 | 94 | 100.0 | 20 | 10 | US-08-457- | Sequence 36, | Applicat | 3.61e-03 |
| 4 | 94 | 100.0 | 20 | 4 | US-07-938- | Sequence 36, | Applicat | 3.61e-03 |
| 5 | 94 | 100.0 | 20 | 10 | US-08-458- | Sequence 36, | Applicat | 3.61e-03 |
| 6 | 94 | 100.0 | 20 | 10 | US-08-457- | Sequence 36, | Applicat | 3.61e-03 |
| 7 | 94 | 100.0 | 20 | 10 | US-08-457- | Sequence 36, | Applicat | 3.61e-03 |
| 8 | 94 | 100.0 | 20 | 8 | US-08-330- | Sequence 62, | Applicat | 3.61e-03 |
| 9 | 94 | 100.0 | 60 | 7 | US-08-226- | Sequence 62, | Applicat | 3.61e-03 |
| 10 | 94 | 100.0 | 60 | 10 | US-08-458- | Sequence 62, | Applicat | 3.61e-03 |
| 11 | 94 | 100.0 | 60 | 10 | US-08-457- | Sequence 62, | Applicat | 3.61e-03 |
| 12 | 94 | 100.0 | 60 | 10 | US-08-457- | Sequence 62, | Applicat | 3.61e-03 |
| 13 | 94 | 100.0 | 60 | 4 | US-07-938- | Sequence 62, | Applicat | 3.61e-03 |
| 14 | 94 | 100.0 | 60 | 10 | US-08-457- | Sequence 62, | Applicat | 3.61e-03 |
| 15 | 94 | 100.0 | 80 | 17 | US-03-112- | Sequence 1, | Applicatio | 3.61e-03 |
| 16 | 94 | 100.0 | 105 | 17 | US-03-142- | Sequence 2, | Applicatio | 3.61e-03 |
| 17 | 94 | 100.0 | 134 | 17 | US-03-142- | Sequence 3, | Applicatio | 3.61e-03 |
| 18 | 94 | 100.0 | 353 | 11 | US-08-526- | Sequence 15, | Applicatio | 3.61e-03 |
| 19 | 94 | 100.0 | 374 | 4 | US-07-938- | Sequence 2, | Applicatio | 3.61e-03 |
| 20 | 94 | 100.0 | 374 | 10 | US-08-457- | Sequence 2, | Applicatio | 3.61e-03 |

ALIGNMENTS

| RESULT | STANDARD | PRT |
|--------|------------------|-------|
| ID | US-08-350-225-36 | 20 AA |

| | | | | | | | | |
|----|----|-------|--------|------------|--------------|-------------|------------|----------|
| 21 | 94 | 100.0 | 374 | 7 | US-08-225- | Sequence 2, | Applicatio | 3 51e-03 |
| 22 | 94 | 100.0 | 374 | 10 | US-08-467- | Sequence 2, | Applicatio | 3 51e-03 |
| 23 | 94 | 100.0 | 374 | 4 | US-07-730- | Sequence 2, | Applicatio | 3 51e-03 |
| 24 | 94 | 100.0 | 374 | 4 | US-07-729- | Sequence 2, | Applicatio | 3 51e-03 |
| 25 | 94 | 100.0 | 374 | 10 | US-08-458- | Sequence 2, | Applicatio | 3 51e-03 |
| 26 | 94 | 100.0 | 374 | 8 | US-08-350- | Sequence 2, | Applicatio | 3 51e-03 |
| 27 | 94 | 100.0 | 374 | 10 | US-08-457- | Sequence 2, | Applicatio | 3 51e-03 |
| 28 | 69 | 73.4 | 20 10 | US-08-467- | Sequence 37, | Applicatio | 5 84e-00 | |
| 29 | 69 | 73.4 | 20 7 | US-08-226- | Sequence 37, | Applicatio | 5 84e-00 | |
| 30 | 69 | 73.4 | 20 8 | US-08-350- | Sequence 37, | Applicatio | 5 84e-00 | |
| 31 | 69 | 73.4 | 20 10 | US-08-457- | Sequence 37, | Applicatio | 5 84e-00 | |
| 32 | 69 | 73.4 | 20 10 | US-08-458- | Sequence 37, | Applicatio | 5 84e-00 | |
| 33 | 69 | 73.4 | 20 10 | US-08-467- | Sequence 37, | Applicatio | 5 84e-00 | |
| 34 | 69 | 73.4 | 20 4 | US-07-538- | Sequence 37, | Applicatio | 1 02e-01 | |
| 35 | 67 | 71.3 | 370 8 | US-08-350- | Sequence 97, | Applicatio | 1 02e-01 | |
| 36 | 67 | 71.3 | 370 10 | US-08-226- | Sequence 97, | Applicatio | 1 02e-01 | |
| 37 | 67 | 71.3 | 370 10 | US-08-457- | Sequence 97, | Applicatio | 1 02e-01 | |
| 38 | 67 | 71.3 | 370 10 | US-08-458- | Sequence 97, | Applicatio | 1 02e-01 | |
| 39 | 67 | 71.3 | 370 10 | US-08-467- | Sequence 97, | Applicatio | 1 02e-01 | |
| 40 | 67 | 71.3 | 370 10 | US-08-467- | Sequence 97, | Applicatio | 1 02e-01 | |
| 41 | 64 | 68.1 | 31 17 | US-09-142- | Sequence 5, | Applicatio | 2 36e-01 | |
| 42 | 64 | 68.1 | 31 17 | US-09-142- | Sequence 4, | Applicatio | 2 36e-01 | |
| 43 | 61 | 64.9 | 367 8 | US-08-150- | Sequence 95, | Applicatio | 5 36e-01 | |
| 44 | 61 | 64.9 | 367 10 | US-08-467- | Sequence 95, | Applicatio | 5 36e-01 | |
| 45 | 61 | 64.9 | 367 10 | US-08-467- | Sequence 95, | Applicatio | 5 36e-01 | |

CC APPLICATION NUMBER: PCT/US93/00139
CC FILING DATE: January 15, 1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Darlene A. Vanstone
CC REGISTRATION NUMBER: 35,729
CC REFERENCE/DOCKET NUMBER: 025.6 US (IM1-228702)
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 466-6040
CC TELEFAX: (617) 466-6040
CC INFORMATION FOR SEQ ID NO: 36:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 20 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC FRAGMENT TYPE: internal
CC SEQUENCE: 20 AA: 2080 MW: 2304 DA:
Query Match: 100.0% Score 94: 15.8 Length 20:
Best Local Similarity 100.0% Pred. No. Mismatches 0: Indels 0:
Matches 13: Conservative 0: Mismatches 0: Indels 0:
DB 8 VFIRKRVSNV:HG 20
QY 1 VFIRKRVSNV:HG 13
RESULT 2
ID US-08-226-248A-J6 STANDARD: PRI: 20 AA
XX XXXXXX
DE Sequence 36, Application US/08226248A
CC Sequence 36, Application US/08226248A
CC GENERAL INFORMATION:
CC APPLICANT: Griffith, Irwin J.
CC APPLICANT: Pollock, Joanne
CC APPLICANT: Bond, Julian F.
CC APPLICANT: Garman, Richard D.
CC APPLICANT: Kuo, Mei-Chang
CC APPLICANT: Yeung, Siu-mei H.
CC APPLICANT: Brauer, Andrew
CC APPLICANT: Exley, Mark A.
CC APPLICANT: Powers, Steven P.
CC TITLE OF INVENTION: Allergic Proteins And Peptides From
CC NUMBER OF SEQUENCES: 20
CC ADDRESS: Immunologic Pharmaceuticals Corporation, Inc
CC STREET: 410 Lincoln St
CC CITY: Waltham
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02154
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/226,248A
CC FILING DATE: April 8, 1994
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/938,990
CC FILING DATE: September 1, 1992
CC APPLICATION NUMBER: PCT/US93/00139
CC FILING DATE: January 15, 1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Darlene A. Vanstone

CC REGISTRATION NUMBER: 35,729
CC REFERENCE/DOCKET NUMBER: 025.6 US (IM1-228CP)
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 466-6000
CC TELEFAX: (617) 466-6040
CC INFORMATION FOR SEQ ID NO: 36:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 20 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC FRAGMENT TYPE: internal
CC SEQUENCE: 20 AA: 2080 MW: 2304 DA:
Query Match: 100.0% Score 94: 15.8 Length 20:
Best Local Similarity 100.0% Pred. No. Mismatches 0: Indels 0:
Matches 13: Conservative 0: Mismatches 0: Indels 0:
DB 8 VFIRKRVSNV:HG 20
QY 1 VFIRKRVSNV:HG 13
RESULT 3
ID US-09-467-006-J6 STANDARD: PRI: 20 AA
XX XXXXXX
DE Sequence 36, Application US/467006
CC Sequence 36, Application US/467006
CC GENERAL INFORMATION:
CC APPLICANT: Griffith, Irwin J.
CC APPLICANT: Pollock, Joanne
CC APPLICANT: Bond, Julian F.
CC APPLICANT: Garman, Richard D.
CC APPLICANT: Kuo, Mei-Chang
CC APPLICANT: Yeung, Siu-mei H.
CC APPLICANT: Brauer, Andrew
CC APPLICANT: Exley, Mark A.
CC APPLICANT: Powers, Steven P.
CC TITLE OF INVENTION: Allergic Proteins And Peptides From
CC NUMBER OF SEQUENCES: 261
CC CORRESPONDENCE ADDRESS:
CC ADDRESS: Immunologic Pharmaceuticals Corporation, Inc
CC STREET: 410 Lincoln St
CC CITY: Waltham
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02154
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/467,006
CC FILING DATE: June 6, 1995
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/350,225
CC FILING DATE: December 6, 1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Jane E. Remillard
CC REGISTRATION NUMBER: 38,872
CC REFERENCE/DOCKET NUMBER: 025.6 USD6 (IM1-028CPD6)
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 227-7400
CC TELEFAX: (617) 227-5941
CC INFORMATION FOR SEQ ID NO: 36:

Best Local Similarity 100.0%; Pred. No. 3.61e-03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 5 VFIKRVSNVILHG 20
II IIIIIIIIIII
QY 1 VFIKRVSNVILHG 13

RESULT 6
ID US-08-467-697-36 STANDARD: PRT: 20 AA.

XX AC xxxxxx

Sequence 36, Application US/08467697

Sequence 36, Application US/08467697

GENERAL INFORMATION:
APPLICANT: Griffith, Irwin J.
APPLICANT: Pollock, Joanne
APPLICANT: Bond, Julian F.
APPLICANT: Garman, Richard D.
APPLICANT: Kuo, Mei-Chang
APPLICANT: Yeung, Siu-mei H.
APPLICANT: Brauer, Andrew
APPLICANT: Exley, Mark A.
APPLICANT: Powers, Steven P.
APPLICANT: Griffith, Irwin J.
TITLE OF INVENTION: Allergenic Proteins And Peptides From
TITLE OF INVENTION: Japanese Cedar Pollen
NUMBER OF SEQUENCES: 261
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
STREET: 610 Lincoln St
CITY: Waltham
STATE: MA
COUNTRY: USA
ZIP: 02154

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION NUMBER: US/08/467,697
FILING DATE: June 5, 1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/350,225
FILING DATE: December 6, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Jane E. Remillard
REGISTRATION NUMBER: 38,872
REFERENCE/DOCKET NUMBER: 025.6 US04 (IM1-028CPD4)
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: Internal
SEQUENCE 20 AA: 2060 MW: 2304 CN:

Query Match 100.0%; Score 94; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.61e-03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 8 VFIKRVSNVILHG 20
IIIIIIIIIIII
QY 1 VFIKRVSNVILHG 13

RESULT 7
ID US-08-467-023-36 STANDARD: PRT: 20 AA.

XX AC xxxxxx

Sequence 36, Application US/08467023

Sequence 36, Application US/08467023

GENERAL INFORMATION:
APPLICANT: Griffith, Irwin J.
APPLICANT: Pollock, Joanne
APPLICANT: Bond, Julian F.
APPLICANT: Garman, Richard D.
APPLICANT: Kuo, Mei-Chang
APPLICANT: Yeung, Siu-mei H.
APPLICANT: Brauer, Andrew
APPLICANT: Exley, Mark A.
APPLICANT: Powers, Steven P.
APPLICANT: Griffith, Irwin J.
TITLE OF INVENTION: Allergenic Proteins And Peptides From
TITLE OF INVENTION: Japanese Cedar Pollen
NUMBER OF SEQUENCES: 261
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
STREET: 610 Lincoln St
CITY: Waltham
STATE: MA
COUNTRY: USA
ZIP: 02154

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,023
FILING DATE: June 6, 1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/350,225
FILING DATE: December 6, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Jane E. Remillard
REGISTRATION NUMBER: 38,872
REFERENCE/DOCKET NUMBER: 025.6 US02 (IM1-028CPD2)
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: Internal
SEQUENCE 20 AA: 2060 MW: 2304 CN:

Query Match 100.0%; Score 94; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.61e-03;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 8 VFIKRVSNVILHG 20
IIIIIIIIIIII
QY 1 VFIKRVSNVILHG 13

RESULT 8
ID US-08-350-225-62 STANDARD: PRT: 50 AA.

XX AC xxxxxx

CC APPLICANT: Young, Siu-mei H.;
 CC APPLICANT: Brauer, Andrew;
 CC APPLICANT: Exley, Mark A.;
 CC APPLICANT: Powers, Steven P.
 CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
 CC TITLE OF INVENTION: Japanese Cedar Pollen
 CC NUMBER OF SEQUENCES: 261
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
 CC STREET: 610 Lincoln St
 CC CITY: Waltham
 CC STATE: MA
 CC COUNTRY: USA
 CC ZIP: 02154
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patent: Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/09/467,697
 CC FILING DATE: June 6, 1995
 CC CLASSIFICATION: 424
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/350,225
 CC FILING DATE: December 6, 1994
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Case E. Remillard
 CC REGISTRATION NUMBER: 38,672
 CC REFERENCE/DOCKET NUMBER: 025.6 USD4 (IMI-025CFD4)
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (617) 227-7430
 CC TELEFAX: (617) 227-5941
 CC INFORMATION FOR SEQ ID NO: 62:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 60 amino acids
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: peptide
 CC FRAGMENT TYPE: internal
 CC SEQUENCE 60 AA: 5644 MW: 19464 CN:
 Query Match 100.0% Score 94: DB 10: Length 60:
 Best Local Similarity 100.0% Pred. No. 3.61e-03:
 Matches 13: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 Db 48 VFIKRVSNVIHG 60
 QY 1 VFIKRVSNVIHG 13
 RESULT 13
 ID US-07-938-920A-62 STANDARD: PRT: 60 AA.
 XX xxxxxx
 DE Sequence 62, Application US/07938990A
 CC Sequence 62, Application US/07938990A
 CC GENERAL INFORMATION:
 CC APPLICANT: Griffith, Irwin J.
 CC APPLICANT: Pollock, Joanne
 CC APPLICANT: Bond, Julian
 CC APPLICANT: Garman, Richard D.
 CC APPLICANT: Kuo, Mei-Chang
 CC APPLICANT: Young, Siu-mei H.
 CC APPLICANT: Brauer, Andrew;
 CC APPLICANT: Exley, Mark A.;
 CC APPLICANT: Powers, Steven P.
 CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
 CC TITLE OF INVENTION: Japanese Cedar Pollen
 CC NUMBER OF SEQUENCES: 70
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Lahive & Cockfield
 CC STREET: Sixty State Street
 CC CITY: Boston
 CC STATE: MA

CC COUNTRY: USA
 CC ZIP: 02109
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC OPERATING SYSTEM: IBM PC compatible
 CC SOFTWARE: Patent: Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: 8/97/938,990A
 CC FILING DATE: 1992890
 CC CLASSIFICATION: 435
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 07/730,462
 CC FILING DATE: July 15, 1991
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 07/729,134
 CC FILING DATE: July 15, 1991
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Amy E. Mandragoras
 CC REGISTRATION NUMBER: 36,207
 CC REFERENCE/DOCKET NUMBER: IPC-025CC (IMI-026)
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (617) 227-7400
 CC TELEFAX: (617) 227-5411
 CC INFORMATION FOR SEQ ID NO: 62:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 60 amino acids
 CC TYPE: AMINO ACID
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: peptide
 CC FRAGMENT TYPE: internal
 CC SEQUENCE 60 AA: 6644 MW: 1464 CN:
 Query Match 100.0% Score 94: DB 4: Length 60:
 Best Local Similarity 100.0% Pred. No. 3.61e-03:
 Matches 13: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 Db 48 VFIKRVSNVIHG 60
 QY 1 VFIKRVSNVIHG 13
 RESULT 14
 ID US-08-467-023-62 STANDARD: PRT: 60 AA.
 XX xxxxxx
 DE Sequence 62, Application US/08467023
 CC Sequence 62, Application US/08467023
 CC GENERAL INFORMATION:
 CC APPLICANT: Griffith, Irwin J.
 CC APPLICANT: Pollock, Joanne;
 CC APPLICANT: Bond, Julia F.;
 CC APPLICANT: Garman, Richard D.
 CC APPLICANT: Kuo, Mei-Chang;
 CC APPLICANT: Young, Siu-mei H.;
 CC APPLICANT: Brauer, Andrew;
 CC APPLICANT: Exley, Mark A.;
 CC APPLICANT: Powers, Steven P.
 CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
 CC TITLE OF INVENTION: Japanese Cedar Pollen
 CC NUMBER OF SEQUENCES: 261
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
 CC STREET: 610 Lincoln St
 CC CITY: Waltham
 CC STATE: MA
 CC COUNTRY: USA
 CC ZIP: 02154
 CC COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/467,023
 FILING DATE: June 6, 1995
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/350,225
 FILING DATE: December 6, 1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Jane E. Remillard
 REGISTRATION NUMBER: 38,872
 REFERENCE/DOCKET NUMBER: 025.5 USJ2 (IMI-02HCPD2)
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 227-7400
 TELEFAX: (617) 227-5941
 INFORMATION FOR SEQ ID NO: 62:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 60 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 FRAGMENT TYPE: internal
 SEQ SEQUENCE 60 AA: 6644 MW: 19464 CN:

Query Match 100.0% Score 94: DB 10: Length 60:
 Best Local Similarity 100.0% Pred. No. 3.61e-03:
 Matches 13: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Db 48 VFIKRVSNVILHG 60
 QY 1 VFIKRVSNVILHG 13

RESULT 15

ID US-09-142-524-1 STANDARD: PRT: 80 AA.
 XX
 AC xxxxxx

Sequence 1, Application US/09142524

Sequence 1, Application US/09142524

GENERAL INFORMATION:
 APPLICANT: Sano, Toshio
 APPLICANT: Kame, Akiohori
 APPLICANT: Katsuki, Kazuo
 APPLICANT: Iwama, Akiko
 APPLICANT: Kano, Kohsuke
 TITLE OF INVENTION: Peptide-based Immunotherapeutic Agent For Treating
 TITLE OF INVENTION: Allergic Diseases
 FILE REFERENCE: Docket No. SPQ-103
 CURRENT APPLICATION NUMBER: US/09/142,524
 CURRENT FILING DATE: 1999-01-04
 EARLIER APPLICATION NUMBER: 8/80/702
 EARLIER FILING DATE: 1996-03-10
 EARLIER APPLICATION NUMBER: PCT/JP97/00740
 EARLIER FILING DATE: 1997-03-10
 NUMBER OF SEQ ID NOS: 5
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 1
 LENGTH: 80
 TYPE: PRT
 ORGANISM: Unknown
 FEATURE:
 OTHER INFORMATION: Description of Unknown Organism: peptide
 SEQ SEQUENCE 80 AA: 9067 MW: 30888 CN:

Query Match 100.0% Score 94: DB 17: Length 80:
 Best Local Similarity 100.0% Pred. No. 3.61e-03:

Matches 13: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 Db 16 VFIKRVSNVILHG 28
 QY 1 VFIKRVSNVILHG 13

Search completed: Mon Jun 19 16:28:35 2000
 Job time : 16 secs.

WATERMAN

(TM)

Release 3.1A John F. Collins, BioComputing Research Unit,
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

WPSrch_pp protein - protein database search, using Smith-Waterman algorithm

Run On: Mon Jun 19 16:17:09 2000; Maspar time 5.14 Seconds

Tabular output not generated. 119,270 Million cell updates/sec

Title: >US-09-142-524A-13
Description: (1-13) from 0509142524A.pep
Perfect Score: 94
Sequence: 1 VF1KRVSNV1HG 13

Scoring table: PAM 150
Gap 15

Searched: 142080 seqs, 47172406 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: pir62

1:pir1 2:pir2 3:pir3 4:pir4

Statistics: Mean 26.420; Variance 35.957; scale 0.735

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description | Pred. No. |
|------------|-------|---------------|--------|-------|-------------|-----------------------|
| 1 | 94 | 100.0 | 374 | 2 | JC2124 | major allergen Cry j |
| 2 | 94 | 100.0 | 374 | 2 | JC2123 | major allergen Cry j |
| 3 | 61 | 64.9 | 326 | 2 | S51509 | pectate lyase - Asper |
| 4 | 57 | 60.6 | 256 | 2 | A32017 | beta-lactamase (EC 3. |
| 5 | 57 | 60.6 | 300 | 2 | F69719 | succinate-CoA ligase |
| 6 | 56 | 59.6 | 257 | 1 | P8BS2S | beta-lactamase (EC 3. |
| 7 | 56 | 59.6 | 257 | 1 | P8BSU2 | beta-lactamase (EC 3. |
| 8 | 55 | 58.5 | 317 | 2 | JC5696 | prolyl aminopeptidase |
| 9 | 55 | 58.5 | 452 | 1 | W2BE4 | gene 4 protein - huma |
| 10 | 55 | 58.5 | 550 | 2 | T15868 | hypothetical protein |
| 11 | 55 | 58.5 | 1894 | 2 | T02155 | DNA-directed DNA poly |
| 12 | 54 | 57.4 | 303 | 2 | F70716 | probable succinyl-coa |
| 13 | 53 | 56.4 | 75 | 1 | S74973 | virulence associated |
| 14 | 53 | 56.4 | 190 | 2 | B23330 | transcription regulat |
| 15 | 53 | 56.4 | 317 | 2 | G70356 | thymidylate synthase |
| 16 | 53 | 56.4 | 375 | 1 | A28630 | muconate cycloisomera |
| 17 | 53 | 56.4 | 475 | 1 | JN0327 | sulfate adenylyltrans |
| 18 | 53 | 56.4 | 835 | 2 | JC6140 | celi surface associat |
| 19 | 53 | 56.4 | 985 | 2 | S99330 | Na+/H-exchanging pro |
| 20 | 52 | 55.3 | 365 | 2 | D75167 | hypothetical protein |
| 21 | 52 | 55.3 | 404 | 2 | S12209 | pectate lyase (EC 4.2 |
| 22 | 52 | 55.3 | 552 | 2 | S46809 | hypothetical protein |
| 23 | 51 | 54.3 | 111 | 2 | A72079 | hypothetical protein |

24 51 54.3 238 2 T16640 hypothetical protein 6.30e-01
25 51 54.3 254 2 E71225 hypothetical protein 6.30e-01
26 51 54.3 261 2 JN0871 ribosomal protein S4 6.30e-01
27 51 54.3 379 2 S56193 probable membrane pro 6.30e-01
28 51 54.3 379 2 S46187 probable membrane pro 6.30e-01
29 51 54.3 381 2 S63322 probable membrane pro 6.30e-01
30 51 54.3 720 2 S50920 hypothetical protein 6.30e-01
31 51 54.3 785 2 B36847 ATP/GTP-binding prote 6.30e-01
32 51 54.3 785 2 D42515 D5R protein - vaccina 6.30e-01
33 51 54.3 785 1 Q0V28 2S protein - vaccinia 6.30e-01
34 51 54.3 810 1 I40508 endopeptidase Clp ATP 6.30e-01
35 51 54.3 819 2 T13355 hypothetical protein 6.30e-01
36 51 54.3 857 1 D35905 endopeptidase Clp ATP 6.30e-01
37 51 54.3 1006 2 C70445 ATPase subunit of ATP 6.30e-01
38 51 54.3 1161 2 B73172 DNA polymerase III, s 6.30e-01
39 50 53.2 267 2 T08079 carbonate dehydratase 9.27e-01
40 50 53.2 267 2 T08082 carbonate dehydratase 9.27e-01
41 50 53.2 270 2 H64248 hypothetical protein 9.27e-01
42 50 53.2 317 2 B74216 endoglucanase - Therm 9.27e-01
43 50 53.2 403 2 JC5462 endopolylacturonase 9.27e-01
44 50 53.2 639 2 F71258 probable heat shock p 9.27e-01
45 50 53.2 661 2 A69252 3-hydroxyacyl-CoA deh 9.27e-01

ALIGNMENTS

1
RESULT JC2124 *type complete
ENTRY major allergen Cry j : precursor (clone pCCI-15) - Japanese
TITLE cedar:
ORGANISM #formal_name Cryptomeria japonica #common_name Japanese cedar
DATE 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change
26-Aug-1999
ACCESSIONS JC2124
REFERENCE JC2123
#authors Sone, T.; Komiya, N.; Shimizu, K.; Kusakabe, T.; Morikubo,
K.; Kino, K.
#journal Biochem. Biophys. Res. Commun. (1994) 199:619-625
#title Cloning and sequencing of cDNA coding for Cry j I, a major
allergen of Japanese cedar pollen.
#cross-references MIM:54183234
#molecule_type mRNA
#residues 1-374 #label SON
#cross-references GB:D26541; NID:9493633; PID:d1006087; PID:q493634
#experimental_source pollen
#note the authors described carbohydrate binding site for:
residue 279
CLASSIFICATION #superfamily for pectate lyase LA159
KEYWORDS glycoprotein; pollen
FEATURE
1-2:
22-374
#domain signal sequence #status predicted #label SIG
#product major allergen Cry j I (clone pCCI-15) #status
predicted #label MATN
158,191,293,354 #binding_site carbohydrate (ASN) (covalent) #status
predicted
SUMMARY #length 374 #molecular-weight 40702 #checksum 3692
Query Match 100.0% Score 94; DB 2; Length 374;
Best Local Similarity 100.0% Pred. No. 6.70e-08;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 129 VF1KRVSNV1HG 141
QY 1 VF1KRVSNV1HG 13
RESULT 2
ENTRY #type complete
TITLE major allergen Cry j I precursor (clone pCCI-2) - Japanese
cedar
ORGANISM #formal_name Cryptomeria japonica #common_name Japanese cedar
DATE 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change


```

#title      The complete genome sequence of the Gram-positive bacterium
             Bacillus subtilis.
#cross-references M1D:98044033
#accession   F69719
#status      preliminary: nucleic acid sequence not shown;
             translation not shown
#molecule_type DNA
#residues    1-300 #label KUN
#cross-references GR 294112; GB:AA009126; NID:92633902;
             PIN:CA313483.1; PID:e1185201; PID:92633982
#experimental_source strain 165

GENETICS
#gene        sscD
#CLASSIFICATION
#KEYWORDS    acid-thiol lyase; ATP; phosphohistidine; phosphoprotein;
             tricarboxylic acid cycle
FEATURE
247         #active_site His (phosphohistidine intermediate) #status
             predicted
SUMMARY      #length 300 #molecular_weight 3392 #checksum 2503
             Query Match      60.6% Score 57; DB 1; Length 300;
             Best Local Similarity 46.2% Pred. No. 5,47e-00;
             Matches 6; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
DB          3 VFINKIRIVVQ 15
            1111111111
CY          1 VFIKRVSNV11G 13

RESULT      6
ENTRY      PNB52S      #type complete
#title     beta-lactamase (EC 3.5.2.6) II precursor - Bacillus sp.
#ALTERNATE_NAMES
#ORGANISM  penicillinase II
#DATE      31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change
             18-Jun-1999
ACCESSIONS A24393; B24393
REFERENCE   Kato, C.; Kudo, T.; Watanabe, K.; Horikoshi, K.
#authors    J. Gen. Microbiol. (1985) 131:3317-3324
#journal    Nucleotide sequence of the beta-lactamase gene of
             alkaliphilic Bacillus sp. strain 170.
#cross-references M1D:86170399
#accession  A24393
#molecule_type DNA
#residues   1-257 #label KAT
#experimental_source strain 170
#accession  B24393
#molecule_type protein
#residues   31-48 #label KAT2
#CLASSIFICATION
#KEYWORDS   antibiotic resistance; hydrolase; metalloprotein; zinc
FEATURE
1-30        #domain signal sequence #status predicted #label SIG\
31-257      #product beta-lactamase II #status predicted\
67          #active_site Glu #status predicted\
116,118,198,240 #binding_site zinc, high affinity (His, His, Cys, His)
             #status experimental
SUMMARY      #length 257 #molecular_weight 28092 #checksum 2889
             Query Match      59.6% Score 56; DB 1; Length 257;
             Best Local Similarity 70.0% Pred. No. 8,33e-00;
             Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
DB          105 FOKRVTDV11 114
            1111111111
CY          2 FIKRVSNV11 11

RESULT      7
ENTRY      PNB5U2      #type complete

```

```

TITLE      beta-lactamase (EC 3.5.2.6) II precursor - Bacillus cereus
ALTERNATE_NAMES
ORGANISM    cephalosporinase II; penicillinase II
#formal_name Bacillus cereus
DATE        04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change
             18-Jun-1999
ACCESSIONS A91806; A91344; A01006
REFERENCE   Hussain, M.; Carlino, A.; Madonna, M.J.; Lampen, J.O.
#authors    J. Bacteriol. (1995) 154:223-229
#journal    Cloning and sequencing of the metallothioester
             beta-lactamase II gene of Bacillus cereus 569/H in
             Escherichia coli.
#cross-references M1D:8508056
#accession  A91806
#molecule_type DNA
#residues   1-257 #label HUS
#cross-references GB:M1185; NID:q142603; PID:AAA22276.1; PID:q142604
#experimental_source strain 569/H
#reference   A91344
#authors    Ambler, R.P.; Daniele, M.; Fleming, J.; Hermoso, C.M.; Pand,
             C.; Waley, S.G.
#journal    FEBS Lett. (1985) 189:207-211
#title     The amino acid sequence of the zinc-requiring beta-lactamase
             II from the bacterium Bacillus cereus 569.
#cross-references M1D:8508445
#accession  A91344
#molecule_type protein
#residues   31-183;187-210;214-257 #label AMB
#experimental_source strain 569/H/H/
COMMENT     beta-lactamase II from Bacillus spp. and beta-lactamase L-1 from
             pseudomonas maltophilia are classified together as class B
             beta-lactamases, which hydrolyze cephalosporins and penicillins.
             beta-lactamase II binds two Zn(II) ions per molecule. Zinc at the
             higher affinity site is necessary for activity of the enzyme.
             Zinc at the lower affinity site (not specified) increases
             hydrolysis of cephalosporin C but not of benzylpenicillin.
CLASSIFICATION #superfamily beta-lactamase II
KEYWORDS       antibiotic resistance; hydrolase; metalloprotein; zinc
FEATURE
1-30          #domain signal sequence #status predicted #label SIG\
31-257        #product beta-lactamase II #status experimental #label
             MPI\
67            #active_site Glu #status predicted\
116,118,198,240 #binding_site zinc, high affinity (His, His, Cys, His)
             #status experimental
SUMMARY      #length 257 #molecular_weight 28092 #checksum 2889
             Query Match      59.6% Score 56; DB 1; Length 257;
             Best Local Similarity 70.0% Pred. No. 8,33e-00;
             Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
DB          105 FOKRVTDV11 114
            1111111111
CY          2 FIKRVSNV11 11

RESULT      8
ENTRY      JC5696      #type complete
#title     prolyl aminopeptidase (EC 3.4.11.5) - Serratia marcescens
#ALTERNATE_NAMES
             cytosol aminopeptidase V; Pro-X aminopeptidase; proline
             aminopeptidase; proline iminopeptidase
ORGANISM      #formal_name Serratia marcescens
DATE          20-Nov-1997 #sequence_revision 20-Nov-1997 #text_change
             29-Sep-1999
ACCESSIONS   JC5696
REFERENCE     JC5696
#authors     Kabashima, T.; Kitazono, A.; Kitano, A.; Ito, K.; Yoshimoto,
             T.
#journal     J. Biochem. (1997) 122:601-605
#title     Prolyl aminopeptidase from Serratia marcescens: Cloning of
             the enzyme gene and crystallization of the expressed
             enzyme.
#cross-references M1D:9806313

```

```

#accession JC5696
#molecule_type DNA
#residues 1-317 #label KAB
#cross-references DDBJ:U87897; NID:q2605614; PDB:BAA23316.1;
#cross-references DDBJ:U87897; NID:q2605615
#comment This enzyme catalyzes the removal of amino-terminal proline from
peptides.

GENETICS
#gene spaf
#classification #superfamily proline aminopeptidase
#keywords alpha-aminocaproylpeptide hydrolase
#summary #length 317 #molecular-weight 35083 #checksum 7818

Query Match 58.5% Score 55: DB 2: Length 317:
Best Local Similarity 41.7% Pred. No. 1,26e-01:
Matches 5: Conservative 0: Mismatches 1: Indels 0: Gaps 0:

ID 254 LRRHIPAVVIRG 265
1 1111111111
2 2 FIKRVSNVLIHG 13

#name WZBE4 #type complete
#description gene 4 protein - human herpesvirus 3
#title #formal_name human herpesvirus 3 varicella-zoster virus
#organism #accession D27212
#date 16-Jul-1999
#authors Davidon, A.J.; Scott, J.E.
#journal J. Gen. Virol. (1986) 57:1789-1816
#title The complete DNA sequence of varicella-zoster virus.
#cross-references MCID:86306557
#accession D27212
#molecule_type DNA
#residues 1-452 #label DAV
#cross-references EMBL:X04370; NID:q59989; PDB:CAA27887.1; PDB:q59993

GENETICS
#gene 4
#classification #superfamily varicella-zoster virus gene 4 protein
#keywords transcription regulation
#summary #length 452 #molecular-weight 51542 #checksum 254

Query Match 58.5% Score 55: DB 1: Length 452:
Best Local Similarity 38.5% Pred. No. 1,26e-01:
Matches 5: Conservative 0: Mismatches 2: Indels 0: Gaps 0:

DB 370 VMIRANTAVVIRG 382
1 1111111111
2 2 FIKRVSNVLIHG 13

#name T15868 #type complete
#description hypothetical protein C56G2.7 - Caenorhabditis elegans
#title #formal_name Caenorhabditis elegans
#organism #accession T15868
#date 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change
20-Sep-1999
#cross-references EMBL:U23177; NID:q726411; PDB:q726419;
PDB:AAA64334.1; CESP:C56G2.7
#experimental_source strain Bristol N2

GENETICS
#gene CESP:C56G2.7

```

```

#introns 13/3: 70/3: 135/1: 165/1: 469/3: 525/1
#summary
#length 550 #molecular-weight 59367 #checksum 1654

Query Match 58.5% Score 55: DB 2: Length 550:
Best Local Similarity 50.0% Pred. No. 1,26e-01:
Matches 6: Conservative 0: Mismatches 1: Indels 0: Gaps 0:

DB 224 VFIKOSNDMLIH 235
1 1111111111
2 2 FIKRVSNVLIHG 12

#name T02155 #type complete
#description DNA-directed DNA polymerase homolog - Arabidopsis thaliana
#title #formal_name Arabidopsis thaliana #common_name mouse ear
#organism #accession T02155
#date 05-Mar-1999 #seq #revision 05-Mar-1999 #text_change
20-Sep-1999
#accessions T02155
#reference 2:4592
#authors Vysotskaya, V.S.; Schwartz, D.R.; Toriumi, M.; Yao, G.; Li,
O.; Kwan, A.; Li, S.; Li, C.; Araceli, R.; Ag, M.; Brendel,
V.; Buehler, E.; Conway, A.B.; Conway, A.B.; Newell, M.;
Feng, J.; Kim, C.; Kurtz, D.; Li, Y.; Palm, C.; Salas,
P.; Sun, H.; Davis, R.W.; Ecker, J.R.; Federspiel, N.A.;
Theologis, A.
#submission submitted to the EMBL Data Library, June 1998
#description Arabidopsis thaliana chromosome 1 BAC T1F15 sequence
#accession T02155
#status preliminary: translated from GB/EMBL/DBS
#molecule_type DNA
#residues 1-1894 #label VVS
#cross-references EMBL:A004393; NID:q318024; PDB:3176691;
GSPB:GN0359; A1SF:1F15 3

GENETICS
#gene
#map_position 1
#introns 38/1: 65/3: 85/1: 201/3: 141/3: 158/3: 181/3: 224/3: 292/1:
321/3: 862/1: 926/1: 1043/3: 1091/2: 1285/1: 1334/3:
1425/3: 1461/3: 1535/2: 1611/3: 1837/3
#summary #length 1894 #molecular-weight 211356 #checksum 9164

Query Match 58.5% Score 55: DB 2: Length 1894:
Best Local Similarity 50.0% Pred. No. 1,26e-01:
Matches 6: Conservative 0: Mismatches 1: Indels 0: Gaps 0:

DB 1693 YAEKRVYVVIHG 1794
1 1111111111
2 2 FIKRVSNVLIHG 13

#name F70716 #type complete
#description probable succinyl-coa synthetase alpha chain - Mycobacterium
tuberculosis (strain H37Rv)
#title #formal_name Mycobacterium tuberculosis
#organism #accession F70716
#date 16-Jul-1999 #sequence_revision 17-Jul-1996 #text_change
16-Jul-1999
#accessions F70716
#reference A70500
#authors Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, J.; Churcher,
C.; Harris, D.; Gordon, S.V.; Eigmeier, K.; Gas, S.; Barry,
III, C.E.; Tekala, F.; Badcock, K.; Basham, D.; Brown, D.;
Chillingworth, T.; Connor, R.; Davies, K.; Devlin, K.;
Fellwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Horsley, T.; Jacobs, K.; Krogh, A.; McLean, J.; Molle, S.;
Murphy, L.; O'Neil, S.; Osborne, J.; Quail, M.A.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.;
Skellton, S.; Squares, R.; Squires, R.; Sulston, J.E.;
Taylor, K.; Whitehead, S.; Barrell, B.G.
#journal Nature (1998) 393:537-544

```

```

#title      Deciphering the biology of Mycobacterium tuberculosis from
#            the complete genome sequence.
#cross-references MUID:98295987
#accession   F7016
##status     preliminary; nucleic acid sequence not shown;
#            translation not shown
##molecule_type DNA
##residues    1-303 ##label CCL
##cross-references GB:279730; GB:AL23456; NID:q13261628;
#            PID:CA801998.1; PID:e264228; PID:g1524210
##experimental_source strain H37Rv

GENETICS
#gene        sucD
#classification #superfamily succinate--CoA ligase (ADP-forming) alpha chain
#summary      #length 333 #molecular-weight 31229 #checksum 2091

Query Match 57.4% Score 54; DB 2; Length 303;
Best Local Similarity 38.5%; Pred. No. 1.90e+01;
Matches 5; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

Db 6 IFLSRDNKVIIVQG 18
Qy 1 VFIRVSNVII 13

RESULT 13
ENTRY S74973 #type complete
#title Virulence associated protein vaps - Synecocystis sp. (strain PCC 6803)
ALTERNATE_NAMES protein ssl2922
ORGANISM #formal_name Synecocystis sp.
#variety PCC 6803
#date 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
ACCESSIONS S74973
REFERENCE S74322
#authors Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.; Hirose, M.; Sugiura, M.; Sasamoto, S.; Kimura, T.; Hosouchi, T.; Matsuno, A.; Muraki, A.; Nakazaki, N.; Naruo, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda, M.; Iabata, S.
#journal DNA Res. (1996) 3:109-136
#title Sequence analysis of the genome of the unicellular cyanobacterium Synecocystis sp. PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions.
#cross-references MUID:97061201
#accession S74973
#status     nucleic acid sequence not shown; translation not shown
##molecule_type DNA
##residues    1-75 ##label KAN
##cross-references EMBL:D90402; GB:AB001339; NID:q1652627;
#            PID:BA017013.1; PID:d1017746; PID:g1652088
##note       the nucleotide sequence was submitted to the EMBL Data Library, June 1996

GENETICS
#gene        vapB
#classification #superfamily virulence-associated protein vaps
#keywords     DNA binding; transcription regulation
#summary      #length 75 #molecular-weight 8539 #checksum 9185

Query Match 55.4% Score 53; DB 1; Length 75;
Best Local Similarity 45.5%; Pred. No. 2.85e+01;
Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 28 VVKKIGSTII 38
Qy 1 VFIRVSNVII 11

RESULT 14
ENTRY B72330 #type complete

```

```

#title      transcription regulator, TetR family - Thermotoga maritima
#            (strain MSB8)
#formal_name Thermotoga maritima
#date 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 20-Aug-1999
ACCESSIONS B72330
REFERENCE A72200
#authors Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.; Dodson, R.J.; Haft, D.; Hickey, E.K.; Peterson, S.D.; Nelson, W.C.; Ketchum, K.A.; McDonald, L.; Utterback, T.R.; Males, J.A.; Linher, A.D.; Garrett, M.M.; Stewart, A.M.; Corton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; Heidelberg, J.; Sutton, G.G.; Fleischmann, R.D.; White, O.; Salzberg, S.L.; Smith, W.O.; Venter, J.C.; Fraser, C.M.
#journal Nature (1999) 399:323-329
#title Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence of Thermotoga maritima
#cross-references MUID:99287316
#accession B72330
##status     preliminary
##molecule_type DNA
##residues    1-190 ##label ARN
##cross-references GB:AE001749; GB:AE00512; NID:q4981365; PID:q4981354;
#            IGR:IM0823
##experimental_source strain MSB8

GENETICS
#gene        TM0823
#summary      #length 190 #molecular-weight 22452 #checksum 1501

Query Match 56.4% Score 53; DB 2; Length 190;
Best Local Similarity 58.3%; Pred. No. 2.85e+01;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Db 126 FIKRVEELIEG 137
Qy 2 FIKRVSNNVII 13

RESULT 15
ENTRY G70356 #type complete
#title Thymidylate synthase complementing protein - Aquifex aeolicus
#formal_name Aquifex aeolicus
#date 08-May-1998 #sequence_revision 08-May-1998 #text_change 08-May-1998
ACCESSIONS G70356
REFERENCE A70356
#authors Derkx, C.; Wilson, P.V.; Bashford, J.; Young, W.G.; Tenckhoff, A.; Zeman, D.B.; Overbeck, R.; Shead, W.A.; Keller, M.; Angli, M.; Butler, K.; Feldman, R.A.; Shuff, J.M.; Olsen, G.J.; Swanson, K.V.
#journal Nature (1998) 392:355-358
#title The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
#cross-references MUID:98195666
#accession G70356
#status     preliminary; nucleic acid sequence not shown;
#            translation not shown
##molecule_type DNA
##residues    1-317 ##label AQF
##cross-references GB:AE000700; NID:q2983248; PID:q2983252; GB:AE000657
##experimental_source strain VF5

GENETICS
#gene        thy
#summary      #length 317 #molecular-weight 37642 #checksum 917

Query Match 56.4% Score 53; DB 2; Length 317;
Best Local Similarity 60.0%; Pred. No. 2.85e+01;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Db 229 VFIRKIEENL 238
Qy 1 VFIRVSNVII 10

```

Search completed: Mon Jun 19 16:17:16 2000
Total time: 7 secs.

 W A R E N
 (TM)

Release 3.1A John F. Collins, BioComputing Research Unit,
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Msearch_pp protein - protein database search, using Smith-Waterman algorithm
 Run on: Mon Jun 19 16:16:17 2000: Masspar time 3.50 Seconds
 Iteration output not generated. 112 992 Million cell updates/sec

Title: >US-09-142-524A-13
 Description: (1-13) from US09142524A.pep
 Perfect Score: 94
 Sequence: : VPIKRVSNVTHG 13
 Scoring table: FAM 15C
 Gap 15

Searched: 33857 seqs, 30454973 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: swiss-prot38
 1:swissprot

Statistics: Mean 27.045; Variance 31.177; scale 0.867

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query Match | Length | ID | Description | Pred. No. |
|------------|-------------|--------|----|-------------|-----------|
| 1 | 94 | 100.0 | 1 | SBP_CHY2A | 7.95e-10 |
| 2 | 61 | 64.9 | 1 | PELLENI | 1.62e-01 |
| 3 | 59 | 62.8 | 1 | L756-CAESL | 4.45e-01 |
| 4 | 57 | 60.6 | 1 | BLAB-BACCE | 1.19e-00 |
| 5 | 57 | 60.6 | 1 | SUCD-BACSU | 1.19e-00 |
| 6 | 56 | 59.6 | 1 | BLA2-BACSP | 1.19e-00 |
| 7 | 56 | 59.6 | 1 | BLA2-BACCE | 1.19e-00 |
| 8 | 56 | 59.6 | 1 | KEPR-NITVU | 1.19e-00 |
| 9 | 55 | 58.5 | 1 | IE63-VZVD | 1.19e-00 |
| 10 | 55 | 58.5 | 1 | YOK7-CAEEL | 3.11e-00 |
| 11 | 54 | 57.4 | 1 | SUCD-MYCTC | 4.98e+00 |
| 12 | 53 | 56.4 | 1 | CATB-PSEPU | 7.91e+00 |
| 13 | 53 | 56.4 | 1 | CYSN-ECOLI | 7.91e+00 |
| 14 | 53 | 56.4 | 1 | NAH1-YEAST | 7.91e+00 |
| 15 | 52 | 55.3 | 1 | 9612-LYCES | 1.25e-01 |
| 16 | 52 | 55.3 | 1 | CPAM-BACSU | 1.25e+01 |
| 17 | 52 | 55.3 | 1 | THN8-YEAST | 1.95e+01 |
| 18 | 51 | 54.3 | 1 | RS4-DROME | 1.95e+01 |
| 19 | 51 | 54.3 | 1 | SCAL-TRIVA | 1.95e+01 |
| 20 | 51 | 54.3 | 1 | SCA3-TRIVA | 1.95e+01 |
| 21 | 51 | 54.3 | 1 | YB8L-YEAST | 1.95e+01 |
| 22 | 51 | 54.3 | 1 | YF02-YEAST | 1.95e-01 |
| 23 | 51 | 54.3 | 1 | YNT6-YEAST | 1.95e-01 |

| | | | | | | |
|----|----|------|------|---|------------|----------|
| 24 | 51 | 54.3 | 601 | 1 | YE06_SCHPO | 1.95e-01 |
| 25 | 51 | 54.3 | 720 | 1 | YM52-YEAST | 1.95e-01 |
| 26 | 51 | 54.3 | 785 | 1 | VD05_VACCC | 1.95e-01 |
| 27 | 51 | 54.3 | 785 | 1 | VD05_VACCV | 1.95e-01 |
| 28 | 51 | 54.3 | 785 | 1 | VD05_VARV | 1.95e-01 |
| 29 | 51 | 54.3 | 810 | 1 | ME8-BACSU | 1.95e-01 |
| 30 | 51 | 54.3 | 819 | 1 | Y083-INPVP | 1.95e-01 |
| 31 | 51 | 54.3 | 857 | 1 | CLPB-ECOLI | 1.95e-01 |
| 32 | 51 | 54.3 | 1161 | 1 | DP1A-BORBU | 1.95e-01 |
| 33 | 50 | 53.2 | 77 | 1 | RUXA-ARCFU | 3.04e-01 |
| 34 | 50 | 53.2 | 129 | 1 | CYT_CYPCA | 3.04e-01 |
| 35 | 50 | 53.2 | 262 | 1 | RS4_CANAL | 3.04e-01 |
| 36 | 50 | 53.2 | 270 | 1 | Y442_MYCCE | 3.04e-01 |
| 37 | 50 | 53.2 | 305 | 1 | K161-RAI | 3.04e-01 |
| 38 | 50 | 53.2 | 587 | 1 | REL-MOUSE | 3.04e-01 |
| 39 | 50 | 53.2 | 633 | 1 | NCQG-RHIS | 3.04e-01 |
| 40 | 50 | 53.2 | 639 | 1 | H1PG-TRIPA | 3.04e-01 |
| 41 | 50 | 53.2 | 788 | 1 | EXGL-CYCCA | 3.04e-01 |
| 42 | 50 | 53.2 | 908 | 1 | H104-YEAST | 3.04e-01 |
| 43 | 50 | 53.2 | 972 | 1 | MIRA-YEAST | 3.04e-01 |
| 44 | 49 | 52.1 | 310 | 1 | PIP-NEICO | 4.68e-01 |
| 45 | 49 | 52.1 | 981 | 1 | GENE_HAEIN | 4.68e-01 |

ALIGNMENTS

RESULT 1
 ID SBP_CHY2A STANDARD; PRI: 374 AA.
 AC P28532;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE SUGI BASIC PROTEIN PRECURSOR (SSP) (NACAR ALLERGEN CRY 3.1) (CRY 3.1).
 OS Cryptomeria japonica (Japanese cedar).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta.
 OC euphyllophytes; Spermatophyta; Coniferopsida; Coniferales;
 OC Taxodiaceae; Cryptomeria.
 RN [1]
 RC SEQUENCE FROM N.A., AND PARTIAL SEQUENCE
 RC TISSUE-POLLEN;
 RX MEDLINE: 94193234.
 RA Sone T., Koniya N., Shimizu K., Kusakabe T., Morikubo K.,
 RA Kino K.;
 RT Cloning and sequencing of cDNA coding for Cry 3.1, a major allergen
 RT of Japanese cedar pollen.
 RL Biochem. Biophys. Res. Commun 199;619-625(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RP TISSUE-POLLEN;
 RA Namba M., Kuroue M., Torioka K., Fukuda S., Kurimoto M.,
 RA Submitted (1992-1994) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RC SEQUENCE OF 22-41.
 RP TISSUE-POLLEN;
 RX MEDLINE: 89031257.
 RA Tanai M., Ando S., Usui M., Kurimoto M., Sakaguchi M., Koike S.,
 RA Matsuda T.;
 RT N-terminal amino acid sequence of a major allergen of Japanese cedar
 RT pollen (Cry 3.1).
 RL FEBS Lett. 239;329-332(1988).
 RN [4]
 RP CARBOHYDRATES
 RX MEDLINE: 95003748.
 RA Hijikata A., Matsumoto I., Kojima K., Ogawa H.;
 RT Antigenicity of the oligosaccharide moiety of the Japanese cedar
 RT (Cryptomeria japonica) pollen allergen, Cry 3.1.
 RL Int. Arch. Allergy Immunol. 105;198-202(1994).
 CC -1- PTM: CONTAINS FUCOSE/XULOSE-CONTAINING N-LINKED OLIGOSACCHARIDES.
 CC -1- DISEASE: THIS PROTEIN IS THE MAJOR ALLERGEN OF JAPANESE CEDAR
 CC POLLEN. THE MOST COMMON POLLEN ALLERGEN IN JAPAN.
 CC -1- MISCELLANEOUS: THE SEQUENCE OF CRY 3.1 FORM A IS SHOWN HERE. FORM
 CC B DIFFERS IN SIX POSITIONS.
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.


```

CC      AMB A I/AMB A II/CRY J I SUBFAMILY.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC
CC      EMBL: D26544; BAA05542.1;
CC      EMBL: D26545; BAA05543.1;
CC      EMBL: D24639; BAA07020.1;
CC      PIR: A44773; A44773;
CC      PFAM: PF00544; pec_lyase;
CC      PRINTS: PR00827; AMBALLERGEN;
CC      Allergen: Glycoprotein; Multigene family; Signal;
CC      SIGNAL 2;
CC      CHAIN 22 374 SUGI BASIC PROTEIN.
CC      VARIANT 12 12 L -> F (IN CRY J I-B);
CC      VARIANT 143 143 H -> Y (IN CRY J I-B);
CC      VARIANT 202 202 S -> T (IN CRY J I-B);
CC      VARIANT 221 221 L -> S (IN CRY J I-B);
CC      VARIANT 358 358 Q -> H (IN CRY J I-B);
CC      VARIANT 361 361 K -> Q (IN CRY J I-B);
CC      VARIANT 368 368 POTENTIAL;
CC      CARBOHYD 148 158 POTENTIAL;
CC      CARBOHYD 191 191 POTENTIAL;
CC      CARBOHYD 243 243 POTENTIAL;
CC      SEQUENCE 374 AA: 46545 MW: 744825950249566F CRC64:
CC
CC      Query Match 100.0% Score 94 DB 1 Length 374;
CC      Best Local Similarity 100.0% Pred. No. 7.99e-10;
CC      Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC      Db 129 VFIRKRVSNV:IRG 141
CC      QY 1 VFIRKRVSNV:IRG 13
CC
CC      RESULT 2
CC      ID PEL-EMENI STANDARD: PRT: 326 AA.
CC      AC Q00645;
CC      DT 15-DEC-1998 (Rel. 37, Created)
CC      DT 15-DEC-1998 (Rel. 37, Last sequence update)
CC      DT 15-DEC-1998 (Rel. 37, Last annotation update)
CC      DE PECTATE LYASE, PRECURSOR (EC 4.2.2.2);
CC      IN PELA.
CC      CS Emericella nidulans (Aspergillus nidulans).
CC      CC Eukaryota; Fungi; Ascomycota; Plecomycetes; Eurotiiales;
CC      Trichomaceae; Emericella.
CC      RN [1]
CC      RP SEQUENCE FROM N.A.
CC      RC STRAIN=FGSC 4;
CC      RX MEDLINE: 95308536
CC      RA Ho M.C., Whitehead M.P., Cleveland T.E., Dean R.A.;
CC      *Sequence analysis of the Aspergillus nidulans pectate lyase pela
CC      gene and evidence for binding of promoter regions to CREA, a
CC      regulator of carbon catabolite repression.*;
CC      Curr. Genet. 27:142-149(1995).
CC      CC -1- CATALYTIC ACTIVITY: ELIMINATIVE CLEAVAGE OF PECTATE TO GIVE
CC      OLIGOSACCHARIDES WITH 4-DEOXY-ALPHA-D-MANN-4-ENURONOSYL GROUPS
CC      AT THEIR NON-REDUCING ENDS
CC      CC -2- SUBCELLULAR LOCATION: SECRETED (POTENTIAL)
CC      CC -3- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC

```

```

DR      EMBL: U05592; AAA80568.1;
DR      PFAM: PF00544; pec_lyase;
DR      KW Lyase; Signal;
DR      FT SIGNAL 1 15 POTENTIAL;
DR      FT CHAIN 16 326 PECTATE LYASE.
DR      SQ SEQUENCE 326 AA: 34580 MW: F6445A4A6D615D49 CRC64:
CC
CC      Query Match 64.9% Score 51 DB 1 Length 326;
CC      Best Local Similarity 63.5% Pred. No. 1.62e-01;
CC      Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
CC
CC      Db 110 YIRRVSNV:IMR 120
CC      QY 2 YIRRVSNV:IMR 12
CC
CC      RESULT 3
CC      ID L756_CAEEL STANDARD: PRT: 413 AA.
CC      AC Q11184;
CC      DT 01-NOV-1997 (Rel. 35, Created)
CC      DT 01-NOV-1997 (Rel. 35, Last sequence update)
CC      DT 15-JUL-1998 (Rel. 36, Last annotation update)
CC      DE LEI-756 PROTEIN.
CC      GN LEI-756 OR COSD11.4.
CC      OS Caenorhabditis elegans.
CC      CC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae.
CC      CC Rhabditina; Rhabditoidea; Rhabditidae; Pelodidae; Caenorhabditis.
CC      RN [1]
CC      RP SEQUENCE FROM N.A.
CC      RC STRAIN=BRISTOL N2;
CC      RA Du 2;
CC      CC -1- SIMILARITY: BELONGS TO THE HEPARIN-BINDING GROWTH FACTORS FAMILY
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC
CC      DR EMBL: J00048; AAB53825.1;
CC      DR HSSP: P05230; 2AXM.
CC      DR WORMPEP; C05D11.4; CE01134.
CC      DR PROSITE; P020247; HGGF_FGF; FALSE_NEG.
CC      DR PFAM: PF00167; FGF; 1.
CC      SQ SEQUENCE 413 AA: 49116 MW: 780070202A00556F CRC64:
CC
CC      Query Match 62.8% Score 59 DB 1 Length 413;
CC      Best Local Similarity 50.0% Pred. No. 4.45e-01;
CC      Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
CC
CC      Db 353 VFVKRNSDV:VIN 364
CC      QY 1 VFVKRNSDV:VIN 12
CC
CC      RESULT 4
CC      ID BLAB_BACCE STANDARD: PRT: 256 AA.
CC      AC P14488;
CC      DT 01-JAN-1990 (Rel. 13, Created)
CC      DT 01-JAN-1990 (Rel. 13, Last sequence update)
CC      DT 15-JUL-1999 (Rel. 38, Last annotation update)
CC      DE BETA-LACTAMASE, TYPE II PRECURSOR (EC 3.5.2.6) (PENICILLINASE)
CC      DE (CEPHALOSPORINASE).
CC      OS Bacillus cereus.
CC      CC Bacteria; Firmicutes; Bacillus/Clostridium group;
CC      CC Bacillus/Staphylococcus group; Bacillus.
CC      RN [1]
CC      RP SEQUENCE FROM N.A.
CC      RC STRAIN=5/B/6;
CC      RX MEDLINE: 88227879.

```


CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch.)
 CC
 CC EMBL: M13350; AAA22276.1; ..
 CC PIR: A24933; PNB525.
 CC HSSP: F04190; 1BMC
 CC PROSITE: PS00743; BETA-LACTAMASE_B_1; 1.
 CC PROSITE: PS00744; BETA-LACTAMASE_B_2; 1.
 CC PFM: PF00753; lactamase_B_1;
 CC Hydrolase; Zinc; Antibiotic resistance; Signal;
 CC SIGNAL 30
 CC CHAIN 31 257 BETA-LACTAMASE, TYPE II;
 CC METAL 116 116 ZINC 1 (BY SIMILARITY);
 CC METAL 118 118 ZINC 1 (BY SIMILARITY);
 CC METAL 120 120 ZINC 2 (BY SIMILARITY);
 CC METAL 179 179 ZINC 1 (BY SIMILARITY);
 CC METAL 198 198 ZINC 2 (BY SIMILARITY);
 CC METAL 240 240 ZINC 2 (BY SIMILARITY);
 CC SEQUENCE 257 AA; 28153 MW; 23EEBDC045AE57; CRC64;

Query Match 59.6% Score 56; DB 1; Length 257;
 Best Local Similarity 70.0%; Pred. No. 1.93e+00;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

1b 105 FOKRVIDVII 114
 1 11111111
 2 FIKRVSNVII 11

RESULT 7
 1b BLA2_NACOE STANDARD. PRT: 257 AA.
 AC PQ4190
 FT 20-MAR-1987 (Rel. 34, Greated)
 FT 20-MAR-1987 (Rel. 34, last sequence update)
 FT 15-FEB-2000 (Rel. 35, last annotation update)
 DE BETA-LACTAMASE, TYPE II PRECURSOR (EC 3.5.2.6) (PENICILLINASE)
 DE (CEPHALOSPORINASE).
 CC BLK:
 CC Bacillus cereus.
 CC Bacteria; Firmicutes; Bacillus/Clostridium group;
 CC Bacillus/Staphylococcus group; Bacillus.
 CC (1)
 CC SEQUENCE FROM N A
 CC STRAIN-569/H;
 CC MEDLINE: 86008056.
 CC Hussain M., Carling A., Mathura K.J., Lamphen J.G.:
 CC "Cloning and sequencing of the metalloprotease beta-lactamase II
 CC gene of *Bacillus cereus* 569/H in *Escherichia coli*.";
 CC *J. Bacteriol.* 164:223-229(1985).
 CC (2)
 CC SEQUENCE OF 31-183; 187-210 AND 214-257.
 CC STRAIN-569/H;
 CC MEDLINE: 86008446.
 CC Ambler R.P., Daniel M., Fleming J., Hermoso J.M., Pang C., Waley S.G.:
 CC "The amino acid sequence of the zinc-requiring beta-lactamase II from
 CC the bacterium *Bacillus cereus* 569.";
 CC *FEBS Lett.* 183:207-211(1985).
 CC (3)
 CC X-RAY CRYSTALLOGRAPHY (3.5 ANGSTROMS).
 CC MEDLINE: 88133841.
 CC Sutton B.J., Artymik P.J., Cordero-Borboa A.E., Little C.,
 CC Phillips D.C., Waley S.G.:
 CC "An X-ray-crystallographic study of beta-lactamase II from *Bacillus*
 CC cereus at 0.35-nm resolution.";
 CC *Biochem. J.* 248:181-188(1987).
 CC (4)
 CC X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).

RX MEDLINE: 96067120.
 RA Carli A., Pares S., Duee E., Galleni M., Duee C., Frere J.-M.,
 RA Dideberg O.:
 RT "The 3-D structure of a zinc metallo-beta-lactamase from *Bacillus*
 RT cereus reveals a new type of protein fold.";
 RL *EMBO J.* 14:4914-4921(1995).
 RN [5]
 RP X-RAY CRYSTALLOGRAPHY (1.85 Å; 357POMS).
 RX MEDLINE: 98437525.
 RA Carli A., Duee E., Galleni M., Frere J.-M., Dideberg O.:
 RT "1.85-Å resolution structure of the zinc (II) beta-lactamase from
 RT *Bacillus cereus*.";
 RL *Acta Crystallogr. D* 54:313-323(1998)
 RN [6]
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
 RX MEDLINE: 98400945.
 RA Fabiane S.M., Sohi M.K., Wan I., Payne D.J., Bateson J.H.,
 RA Mitchell T., Sutton B.J.:
 RT "Crystal structure of the zinc-dependent beta-lactamase from *Bacillus*
 RT cereus at 1.9-Å resolution: binuclear active site with features of a
 RT mononuclear enzyme.";
 RL *Biochemistry* 37:12404-12411(1998).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (1.7 Å; 37OMS).
 RX MEDLINE: 98400945.
 RA Fabiane S.M., Sohi M.K., Sutton B.J.:
 RT Submitted (SEP-1997) to the PDB data bank.
 CC FUNCTION: CAN HYDROLYZE CARBAPENEM COMPOUNDS.
 CC CATALYTIC ACTIVITY: A BETA-LACTAM - H(2)O - A SUBSTITUTED BETA-
 CC AMINO ACID.
 CC COFACTOR: BINDS TWO ZINC IONS PER MOLECULE. THE ENZYME CAN ALSO
 CC FUNCTION WITH ONLY ONE ZINC ION.
 CC SUBUNIT: MONOMER.
 CC SIMILARITY: BELONGS TO THE CLASS-B BETA-LACTAMASE FAMILY

This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch.)
 CC
 CC EMBL: M1189; AAA22276.1; ..
 CC PIR: A01006; PNB502.
 CC PDB: 1BMC; 28-AUG-96.
 CC PDB: 1BME; 07-JUL-97.
 CC PDB: 1BVT; 23-SEP-99.
 CC PDB: 1BC2; 14-OCT-99.
 CC PDB: 1BZ2; 20-APR-99.
 CC PDB: 1BC2; 20-APR-99.
 CC PROSITE: PS00743; BETA-LACTAMASE_B_1; 1.
 CC PROSITE: PS00744; BETA-LACTAMASE_B_2; 1.
 CC PFM: PF00753; lactamase_B_1;
 CC Hydrolase; Zinc; Antibiotic resistance; Signal; 3D-struct.res.
 CC SIGNAL 30
 CC CHAIN 31 257 BETA-LACTAMASE, TYPE II;
 CC METAL 116 116 ZINC 1, HIGH-AFFINITY;
 CC METAL 118 118 ZINC 1, HIGH-AFFINITY;
 CC METAL 120 120 ZINC 2, LOW-AFFINITY;
 CC METAL 179 179 ZINC 2, HIGH-AFFINITY;
 CC METAL 198 198 ZINC 2, LOW-AFFINITY;
 CC METAL 240 240 ZINC 2, LOW-AFFINITY;
 CC SEQUENCE 257 AA; 28092 MW; 268EBF7DDA45431; CRC64;

Query Match 59.6% Score 56; DB 1; Length 257;
 Best Local Similarity 70.0%; Pred. No. 1.93e+00;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
 Db 105 FOKRVIDVII 114
 1 11111111
 2 FIKRVSNVII 11
 QY 2 FIKRVSNVII 11

```

RESULT 8
ID KPRR_NITVU STANDARD: PRT: 290 AA.
AC P37100;
DI 01-OCT-1994 (Rel. 30, Created)
DI 01-OCT-1994 (Rel. 30, Last sequence update)
DI 01-OCT-1994 (Rel. 30, Last annotation update)
DE PHOSPHORIBULOKINASE (EC 2.7.1.19) (PHOSPHOPENTOKINASE) (PRK).
GN CBPP.
OS Nitro bacter vulgaris.
OC Bacteria: Proteobacteria: alpha subdivision: Bradyrhizobium group;
OC Nitrobacteraceae: Nitrobacter.
RN [1]
SE SEQUENCE FROM N.A.
RC STRAIN-T3;
RA Strecker M., Sickinger E., English R.S., Shively J.M., Bock E.;
RL Submitted (XX-1994) to the EMBL/GenBank/DBJ databases.
CC -!- CATALYTIC ACTIVITY: ATP + D-RIBULOSE 5-PHOSPHATE -> ADP +
CC D-RIBULOSE 1,5-BISPHOSPHATE.
CC -!- PATHWAY: CALVIN CYCLE.
CC -!- SIMILARITY: BELONGS TO THE PHOSPHORIBULOKINASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: L22884; AAA25506.1;
CC HSSP: P12033; 1A7J.
CC PR-NTS: PRC0478; PHRIBULOKINASE.
CC PROSITE: PS00567; PHOSPHORIBULOKINASE; 1.
CC PFAM: PF00485; PRK; 1.
KW Transferase; Kinase; Calvin cycle; ATP-binding; Photosynthesis.
FT NP_31925; 12
SC SEQUENCE 290 AA; 34908 MW; 9E297231942541C CRC64;

Query Match 59.6%; Score 55; DB 1; Length 240;
Best Local Similarity 41.7%; Pred. No. 1.93e-00;
Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

DB 254 PMSRANSVING 265
QY 2 FIKRVSNIH 13

RESULT 9
ID IE61_VZVD STANDARD: PRT: 452 AA.
AC P09259;
DI 01-MAR-1995 (Rel. 10, Created)
DI 01-MAR-1995 (Rel. 10, Last sequence update)
DI 01-APR-1993 (Rel. 25, Last annotation update)
DE TRANSCRIPTIONAL REGULATOR IE63 HOMOLOG.
GN 4.
OS Varicella-zoster virus (strain Dumas) (VZV).
OC Viruses: dsDNA viruses, no RNA stage: Herpesviridae;
OC Alphaherpesvirinae: Varicelloviruses.
RN [1]
SE SEQUENCE FROM N.A.
RA MEDLINE: 86306657.
RX Davison A.J., Scott J.E.;
RT "The complete DNA sequence of varicella-zoster virus."
RL J. Gen. Virol. 67:1759-1816(1986).
CC -!- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL54,
CC HSV-2 UL54, EBV-1 5, VZV 4, EBV BMLF1, HCMV UL65, AND HVS-1 57.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X04370; CAA27887.1;
CC PIR: D27212; W2BE4.
CC TRANSCRIPTION regulation.
SQ SEQUENCE 452 AA; 51543 MW; 42926E4A71E380B4 CRC64;

Query Match 58.5%; Score 55; DB 1; Length 452;
Best Local Similarity 38.5%; Pred. No. 3.11e-00;
Matches 5; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

DB 370 VMIARIANIVVRG 382
QY 1 VFIKRVSNVING 13

RESULT 10
ID YQK7_CAEEL STANDARD: PRT: 550 AA.
AC Q09289;
DI 01-NOV-1997 (Rel. 35, Created)
DI 01-NOV-1997 (Rel. 35, Last sequence update)
DI 01-NOV-1997 (Rel. 35, Last annotation update)
DE HYPOTHEICAL 59.4 KD PROTEIN C56G2.7 IN CHROMOSOME 111.
GN C56G2.7.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae.
OC Rhabditia; Rhabditidae; Rhabditidae; Peleodidae; Caenorhabditis
RN [1]
SE SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Connell M.;
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO HUMAN 110 KD CELL MEMBRANE GLYCOPROTEIN
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: U23177; AAA64334.1;
CC WORMPEP: C56G2.7; CE01876.
CC Hypothetical protein.
SQ SEQUENCE 550 AA; 59367 MW; B7E847A58EF0AEE8 CRC64;

Query Match 58.5%; Score 55; DB 1; Length 550;
Best Local Similarity 50.0%; Pred. No. 3.11e-00;
Matches 5; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

DB 224 VFIKRSDNDLH 235
QY 1 VFIKRVSNVING 12

RESULT 11
ID SUCC_MYCTU STANDARD: PRT: 303 AA.
AC P71558;
DI 01-NOV-1997 (Rel. 35, Created)
DI 01-NOV-1997 (Rel. 35, Last sequence update)
DI 15-FEB-2000 (Rel. 39, Last annotation update)
DE SUCCINYL-CCA SYNTHETASE ALP1 CHAIN (EC 6.2.1.5) (SCS-ALPHA).
GN SUCC OR RV0952 OR MTCY10D7.22C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Mycobacterium
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium
RN [1]
SE SEQUENCE FROM N.A.
RC STRAIN-H37RV;
RX MEDLINE: 98235987.
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,

```

RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies K., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jacobs K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne K., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E.,
 RA Taylor K., Whitehead S., Barrett B.G.,
 RA "Deciphering the Biology of Mycobacterium tuberculosis from the
 RA complete genome sequence."
 RL Nature 393:537-544(1998).
 CC -!- CATALYTIC ACTIVITY: SUCCINATE + COA + ATP -> SUCCINYL-COA + ADP +
 CC ORTHOPHOSPHATE.
 CC -!- SUBUNIT: Tetramer of two alpha and two beta chains
 CC (BY SIMILARITY)
 CC -!- SIMILARITY: TO OTHER ALPHA SUBUNITS OF SUCCINYL-COA SYNTHETASE.
 CC OF MALATE-COA LIGASE AND TO A.P. CITRATE-LIGASE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch.)
 CC -----
 CC EMBL: Z79700; CAB01959.1;
 CC RSSP: PC7459; ISCU.
 CC PROSITE: PS01216; SUCCINYL_COA_LIG_1; 1.
 CC PROSITE: PS01499; SUCCINYL_COA_LIG_2; 1.
 CC PFAM: PF00549; ligase-coa; 1.
 CC TUBERCULOSIS: K00952; ..
 KW Ligase; Tricarboxylic acid cycle; ATP-binding; Phosphorylation.
 FT ACT_SITE 259 PHOSPHORYLATED IN THE COURSE OF
 FT CATALYSIS (BY SIMILARITY).
 SQ SEQUENCE 303 AA; 31229 MW; 57769CE8DC0CB48 CRC64;
 Query Match 57.4%; Score 54; DB 1; Length 303;
 Best Local Similarity 38.5%; Pred. No. 4; 96e+00;
 Matches 5; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
 DR 6 IFLSKDKVKVVG 18
 DY 1 VFIKRVSNVTHS 13
 RESULT 12
 ID CATB_PSEP0 STANDARD: PRT: 374 AA.
 AC P08310;
 DT 01-AUG-1988 (Rel. 08; Created);
 DT 01-MAY-1991 (Rel. 18; Last sequence update);
 DT 15-JUL-1999 (Rel. 38; Last annotation update);
 DE MUONATE CYCLOISOMERASE I (EC 5.5.1.1) (CIS,CIS-MUONATE LACTONIZING
 DE ENZYME I) (MIE);
 GN CATB.
 OS Pseudomonas putida.
 SC Bacteria; Proteobacteria; gamma subdivision; Pseudomonas group;
 CC Pseudomonas.
 CC (1)
 RN SEQUENCE FROM N.A.
 RP STRAIN-PRS2000;
 RX MEDLINE: 97277391.
 RA Aldrich T.L., Frantz B., Gill J.F., Kibane J.J., Chakrabarty A.M.;
 RA "Cloning and complete nucleotide sequence determination of the catB
 RA gene encoding cis,cis-muonate lactonizing enzyme."
 RL Gene 52:185-195(1987).
 RN (2)
 RN SEQUENCE FROM N.A.
 RX MEDLINE: 88139192.
 RA Aldrich T.L., Chakrabarty A.M.;
 RA "Transcriptional regulation, nucleotide sequence, and localization of
 RA the promoter of the catB operon in Pseudomonas putida";
 RL J. Bacteriol. 170:1297-1304(1988).
 RN (3)

RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
 RX MEDLINE: 87283908.
 RA Goldman A., Ollis D.L., Steitz T.A.;
 RA "Crystal structure of muonate lactonizing enzyme at 3-A resolution".
 RL J. Mol. Biol. 194:143-153(1997).
 RN (4)
 RP X-RAY CRYSTALLOGRAPHY (1.85 ANGSTROMS).
 RX STRAIN-PRS2000;
 RX MEDLINE: 95107379.
 RA Helein S., Kahn P.C., Guha B., Mallows D.J., Goldman A.;
 RA "The refined x-ray structure of muonate lactonizing enzyme from
 RA Pseudomonas putida PRS2000 at 1.85-A resolution".
 RL J. Mol. Biol. 254:918-942(1995).
 RN (5)
 RP SIMILARITY TO MR.
 RX MEDLINE: 91015392.
 RA Neidhart D.J., Kenyon G.L., Gerit J.A., Petsko G.A.;
 RA "Mandelate racemase and muonate lactonizing enzyme are
 RA mechanistically distinct and structurally homologous."
 RL Nature 347:692-694(1990).
 CC -!- FUNCTION: CATALYZES A SYN CYCLOISOMERIZATION.
 CC -!- CATALYTIC ACTIVITY: 2,5-DIHYDRO-5-OXOPURAN-2-ACETATE -> CIS,CIS-
 CC HEXADIENEDIACETATE.
 CC -!- COFACTOR: REQUIRES MANGANESE.
 CC -!- PATHWAY: SECOND STEP IN THE DEGRADATION OF CATECHOL TO SUCCINATE
 CC AND ACETYL-COA IN THE BETA-KETOACIDATE PATHWAY.
 CC -!- SUBUNIT: HOMODIMER.
 CC -!- SIMILARITY: BELONGS TO THE MANDELATE RACEMASE / MUONATE
 CC LACTONIZING ENZYME FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch.)
 CC -----
 CC EMBL: M16560; AAA25765.1;
 CC EMBL: M16336; AAA25765.1;
 CC PIR: A27316; A27316.
 CC PIR: A28530; A28630.
 CC PDB: 1MUC; 11-JUL-96
 CC PDB: 2MUC; 04-NOV-98.
 CC PDB: 3MUC; 04-NOV-98.
 CC PROSITE: PS00908; MR_MIE_1; 1.
 CC PROSITE: PS00909; MR_MIE_2; 1.
 CC PFAM: PF01188; MR_MIE_1; 1.
 CC Aromatic hydrocarbons catabolism; 13-methylase; Manganese; 13-structure.
 FT ACT_SITE 170 170
 FT METAL 199 199 MANGANESE
 FT METAL 225 225 MANGANESE
 FT METAL 250 250 MANGANESE
 FT CONFLICT 32 32 I -> S (IN REF. 2).
 FT CONFLICT 64 64 S -> I (IN REF. 2).
 FT CONFLICT 243 243 MISSING (IN REF. 2).
 SQ SEQUENCE 374 AA; 40378 MW; A7767A865EACDD8C CRC64;
 Query Match 56.4%; Score 53; DB 1; Length 374;
 Best Local Similarity 41.7%; Pred. No. 7; 91e+00;
 Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 DB 3 VLIETIEAIVH 14
 DY 1 VFIKRVSNVTHS 12
 RESULT 13
 ID CYSN_ECOLI STANDARD: PRT: 475 AA.
 AC P23845;
 DT 01-NOV-1991 (Rel. 20; Created)
 DT 01-NOV-1991 (Rel. 20; Last sequence update)

15-DEC-1998 (rel. 37, Last annotation update)
 DE SULFATE ADENYLATE TRANSFERASE SUBUNIT 1 (EC 2.7.7.4) (ATP-
 DE SULFURYLASE).
 GN CYSN.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-5.
 RC STRAIN-K12:
 RA MEDLINE: 92426517.
 RX Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner F.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.,
 R1 "The complete genome sequence of Escherichia coli K-12."
 R1 Science 277:1453-1474 (1997).
 R1 [3]
 RN CHARACTERIZATION:
 RP MEDLINE: 89115389.
 RX Levh T.S., Taylor J.C., Markham G.D.,
 RA "The sulfate activation locus of Escherichia coli K12: cloning,
 R1 genetic, and enzymatic characterization."
 R1 J. Biol. Chem. 263:2409-2416 (1988).
 RL -!- FUNCTION: MAY BE THE GTPASE, REGULATING ATP SULFURYLASE ACTIVITY.
 CC -!- CATALYTIC ACTIVITY: ATP + SULFATE = PYROPHOSPHATE +
 CC ADENYLSULFATE.
 CC -!- ENZYME REGULATION: GTPASE ACTIVITY IS COUPLED TO STIMULATION OF
 CC THE RATE OF APS FORMATION.
 CC -!- PATHWAY: FIRST STEP IN THE SULFATE ACTIVATION PATHWAY. THIS
 CC REACTION OCCURS EARLY IN THE REDUCTIVE BRANCH OF THE CYSTEINE
 CC BIOSYNTHETIC PATHWAY.
 CC -!- SUBUNIT: HETERODIMER COMPOSED OF CYSD, THE SMALLER SUBUNIT, AND
 CC CYSN.
 CC -!- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
 CC CYSN/ND00 SUBFAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: M74596; AAA23646.1;
 DR EMBL: U29579; AAA69281.1;
 DR EMBL: A5000358; AAC75793.1;
 DR PIR: JN0327; JN0327.
 DR HSP: P02990; LETU.
 DR ECO2DBASE: C058.3; 6TH EDITION.
 DR ECOGENE: EG:0194; CYSN.
 DR PRINTS: PR00315; ELONGATNFT.
 DR PROSITE: PS00301; EFATOR_GTP: 1.
 DR PFAM: PF00009; GTP_EFTU: 1.
 KW Cysteine biosynthesis; Transferase; Nucleotidyltransferase;
 KW GTP-binding.
 FT NP_BIND 34 41 GTP (BY SIMILARITY).
 FT NP_BIND 113 117 GTP (BY SIMILARITY).
 FT NP_BIND 169 171 GTP (BY SIMILARITY).
 CC SEQUENCE 475 AA: 52558 MW: 915983649732A15 CRC64:

Query Match 56.4% Score 53; DB 1; Length 475;
 Best Local Similarity 54.5% Pred. No. 7.91e+00;
 Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

DB 419 IFIDRLSNVTY 429
 QY 1 VFIRKRVSNVII 11
 RESULT 14
 ID NAHL YEAST STANDARD: PRT: 985 AA.
 AC 099271.
 DI 01-NOV-1997 (rel. 35, Created)
 DT 01-NOV-1997 (rel. 35, Last sequence update)
 DT 15-FEB-2000 (rel. 39, Last annotation update)
 DE NA(+)H(+) ANTIporter.
 GN NHA1 OR YLR138W OR L3149 OR L9606.4
 OS Saccharomyces cerevisiae "baker's yeast"
 OC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
 OC Saccharomycetaceae; Saccha. myces.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C / AB972:
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
 RA Favallo A., Fulton L., Gattung S., Greco I., Kirsten J., Kucaba T.,
 RA Hallsworth K., Hawkins J., Hillier L., Jier M., Johnson C.,
 RA Johnston L., Langston T., Latreille P., Le T., Mardis E., Menezes S.,
 RA Miller N., Nhan M., Pauley A., Peluso D., Rifken L., Riles L.,
 RA Taich A., Trevisan E., Vignati D., Wilcox L., Wohldman P., Yeudin M.,
 RA Wilson R., Waterston R.;
 RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-S288C:
 RL Delius H., Hebling U.;
 RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 145-985 FROM N.A.
 RA Rieger M., Mueller-Auer S., Brueckner M.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases
 CC -!- FUNCTION: SODIUM EXPORT FROM CELL. TAKES UP EXTERNAL PROTONS IN
 CC EXCHANGE FOR INTERNAL SODIUM IONS (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
 CC -!- SIMILARITY: BELONGS TO THE FUNGAL NA(+)/H(+) EXCHANGER FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: U53841; AAA82392.1;
 DR EMBL: X91259; CAA52653.1;
 DR EMBL: Z73310; CAA37709.1;
 DR EMBL: Z73311; CAA97711.1;
 DR SCS: L0004113; NHA1.
 KW Transmembrane; Glycoprotein; Sodium transport; Transport; Symport.
 FT TRANSMEM 13 33 POTENTIAL.
 FT TRANSMEM 37 57 POTENTIAL.
 FT TRANSMEM 71 91 POTENTIAL.
 FT TRANSMEM 106 126 POTENTIAL.
 FT TRANSMEM 129 149 POTENTIAL.
 FT TRANSMEM 177 197 POTENTIAL.
 FT TRANSMEM 204 224 POTENTIAL.
 FT TRANSMEM 245 265 POTENTIAL.
 FT TRANSMEM 295 315 POTENTIAL.
 FT TRANSMEM 320 340 POTENTIAL.
 FT TRANSMEM 362 382 POTENTIAL.
 FT TRANSMEM 411 431 POTENTIAL.
 CC SEQUENCE 985 AA: 109369 MW: AD0B86D483DCDD89 CRC54:

Query Match 56.4% Score 53; DB 1; Length 985;
 Best Local Similarity 36.4% Pred. No. 7.91e+00;
 Matches 4; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

Search completed: Mon Jun 19 16:16:23 2000
Job time : 6 secs.

WATERBURY
***** (TM)

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, G.K.
Distribution rights by Oxford Molecular Ltd

Wfsrcch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Mon Jun 19 16:16:40 2000; Maspar time 8.57 Seconds
Tabular output not generated. 105.118 Million cell updates/sec

Title: >US-09-142-524A-13
Description: (1-13) from US09142524A.pep
Perfect Score: 94
Sequence: 1 VF:KPVSNVNIHG 13

Scoring table: PAM 150
Gap 15

Searched: 225878 seqs, 69334122 residues
Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: spiremb12
1:sp-archaea 2:sp-bacteria 3:sp-fungi 4:sp-human
5:sp-invertebrate 6:sp-mammal 7:sp-mhc 8:sp-organelle
9:sp-phage 10:sp-plant 11:prokaryote 12:sp-unclassified
13:sp-vertebrate 14:sp-virus

Statistics: Mean 26.111; Variance 31.885; scale 0.919

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

| SUMMARIES | | | | SUMMARIES | | | |
|------------|-------|--------------|--------|-------------------------|-----------|------------|-------|
| Result No. | Score | Match length | DB ID | Description | Pred. No. | Result No. | Score |
| 1 | 64 | 68.1 | Q9W15 | ORF MSV077 HYPOTHETICAL | 9.02e-02 | 1 | 64 |
| 2 | 63 | 67.0 | Q9ZGL9 | HYPOTHETICAL 41.5 KD P | 1.50e-01 | 2 | 63 |
| 3 | 61 | 64.9 | Q9ZNU7 | POLLEN MAJOR ALLERGEN | 4.07e-01 | 3 | 61 |
| 4 | 61 | 64.9 | Q96385 | CHAGI PRECURSOR | 4.07e-01 | 4 | 61 |
| 5 | 59 | 62.8 | Q65388 | FL121.22 PROTEIN | 1.08e+00 | 5 | 59 |
| 6 | 58 | 61.7 | Q9ZK39 | BETA-N-ACETYLGLUCOSAM | 1.76e+00 | 6 | 58 |
| 7 | 56 | 59.6 | Q9ZK39 | GLYCOSYLTRANSFERASE GT | 4.52e+00 | 7 | 56 |
| 8 | 56 | 59.6 | Q9ZK39 | POTASSIUM FLIPPASE | 4.52e+00 | 8 | 56 |
| 9 | 56 | 59.6 | Q9ZK39 | KO7A12.2 PROTEIN | 4.52e+00 | 9 | 56 |
| 10 | 55 | 58.5 | Q32449 | PROLYL AMINOPEPTIDASE | 7.20e+00 | 10 | 55 |
| 11 | 55 | 58.5 | Q64795 | TIF15.3 PROTEIN (EC 2 | 7.20e+00 | 11 | 55 |
| 12 | 54 | 57.4 | Q82690 | MAT1 | 1.14e+01 | 12 | 54 |
| 13 | 53 | 56.4 | Q72994 | VIRULENCE ASSOCIATED P | 1.79e+01 | 13 | 53 |
| 14 | 53 | 56.4 | Q9ZK70 | TRANSCRIPTIONAL REGULA | 1.79e+01 | 14 | 53 |
| 15 | 53 | 56.4 | Q66883 | THYMIDYLATE SYNTHASE C | 1.79e+01 | 15 | 53 |
| 16 | 53 | 56.4 | Q9ZK70 | CODED FOR BY C. ELEGAN | 1.79e+01 | 16 | 53 |
| 17 | 53 | 56.4 | Q9W11 | HYPOTHETICAL 41.2 KD P | 1.79e+01 | 17 | 53 |
| 18 | 53 | 56.4 | Q22370 | PTTA | 1.79e+01 | 18 | 53 |
| 19 | 53 | 56.4 | Q46322 | PLASMOD GR0497C, COMPL | 1.79e+01 | 19 | 53 |
| 20 | 53 | 56.4 | P90755 | C27A7.3 PROTEIN | 1.79e+01 | 20 | 53 |

| | | | | | | | |
|----|----|------|------|----|---------|------------------------|----------|
| 21 | 53 | 56.4 | 835 | 2 | Q56012 | CELL SURFACE-ASSOCIATE | 1.79e-01 |
| 22 | 53 | 56.4 | 871 | 3 | Q74833 | DNA LIGASE PROTEIN | 1.79e-01 |
| 23 | 52 | 55.3 | 192 | 14 | Q81434 | ENVELOPE PROTEIN E1 F | 2.79e-01 |
| 24 | 52 | 55.3 | 394 | 10 | Q65457 | PECTATE LYASE LIKE PRO | 2.79e-01 |
| 25 | 52 | 55.3 | 394 | 10 | Q65456 | PECTATE LYASE LIKE PRO | 2.79e-01 |
| 26 | 52 | 55.3 | 405 | 10 | Q24216 | PECTATE LYASE | 2.79e-01 |
| 27 | 52 | 55.3 | 427 | 10 | Q19815 | FI7A14.4 PROTEIN | 2.79e-01 |
| 28 | 51 | 54.3 | 111 | 2 | C 2882 | HYPOTHETICAL 12.2 KD P | 4.33e-01 |
| 29 | 51 | 54.3 | 186 | 2 | C 1109 | HYPOTHETICAL 20.9 KD P | 4.33e-01 |
| 30 | 51 | 54.3 | 238 | 5 | C 1411 | COSMID PDB1 | 4.33e-01 |
| 31 | 51 | 54.3 | 254 | 1 | C 1799 | 254AA LONG HYPOTHETICA | 4.33e-01 |
| 32 | 51 | 54.3 | 309 | 5 | C 15568 | ADHESIN PROTEIN AP33-1 | 4.33e-01 |
| 33 | 51 | 54.3 | 309 | 5 | C 15569 | ADHESIN PROTEIN AP33-3 | 4.33e-01 |
| 34 | 51 | 54.3 | 347 | 5 | Q97246 | PC00645W PROTEIN | 4.33e-01 |
| 35 | 51 | 54.3 | 365 | 5 | Q7241 | HEXAMERIN (FRAGMENT) | 4.33e-01 |
| 36 | 51 | 54.3 | 417 | 5 | C 18093 | FI1H7.10 PROTEIN | 4.33e-01 |
| 37 | 51 | 54.3 | 581 | 5 | C 22977 | COSMID F45F2 | 4.33e-01 |
| 38 | 51 | 54.3 | 785 | 14 | Q93119 | ROTATIVE 90.4K PROTEIN | 4.33e-01 |
| 39 | 51 | 54.3 | 1006 | 2 | C 67588 | ATPASE SUBUNIT OF ATP | 4.33e-01 |
| 40 | 50 | 53.2 | 157 | 2 | C 54168 | PEPTIDYDROLYL ISOMERA | 5.68e-01 |
| 41 | 50 | 53.2 | 225 | 14 | Q24989 | ENVELOPE POLYPROTEIN | 5.68e-01 |
| 42 | 50 | 53.2 | 260 | 1 | C 1009 | CONSERVED HYPOTHETICAL | 5.68e-01 |
| 43 | 50 | 53.2 | 286 | 5 | C 1047 | H1ZC20.4 PROTEIN | 5.68e-01 |
| 44 | 50 | 53.2 | 439 | 5 | C 6197 | REVERSE TRANSCRIPTASE | 5.68e-01 |
| 45 | 50 | 53.2 | 450 | 10 | C 10319 | PECTATE LYASE HOMOLOG | 5.68e-01 |

ALIGNMENTS

RESULT ID Q9W15 PRELIMINARY: PRT: 598 AA.
AC Q9W15:
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE ORF MSV077 HYPOTHETICAL PRO1 N.
GN MSV077.
OS Melanoplus sanguinipes entomopoxvirus.
CC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Entomopoxvirinae;
OC Entomopoxvirus B.
PP [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TUCSON.
RX MEDLINE: 9910262.
RA AFONSO C.L., TULMAN E.R., LU Z., OMA E., KUTISH G.F., ROCK D.L.
RI The genome of Melanoplus sanguinipes entomopoxvirus.
RL J. Virol. 73:533-552(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-TUCSON.
RA AFONSO C.L., TULMAN E.R., LU Z., OMA E., KUTISH G.F., ROCK D.L.
RI Submitted (MAY-1998) to the EMBL/GenBank/CDR databases.
RL EMBL: AF083866; AA:976321.
SQ SEQUENCE 598 AA: 73312 MW: 14428F93 CRC32:

Query Match 68.1%; Score 64; DB 14; Length 598.
Best Local Similarity 61.5%; Pred. No. 9.02e-02;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

DB 105 VKRNVSIILIDG 117
QY 1 VKRNVSVNIHG 13

RESULT ID Q9ZGL9 PRELIMINARY: PRT: 360 AA.
AC Q9ZGL9:
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE HYPOTHETICAL 41.5 KD PROTEIN
OS Leptospira borgpetersenii
OC Bacteria; Spirochaetales; Leptospiraceae; Leptospira.


```

FN 123
RP SEQUENCE FROM N.A.
RC STRAIN=L171:
KA KAJAMBAHETI T., RULACH D.M., RAJAKUMAR K., ADLER B.:
RT "Genetic Organization of the Lipopolysaccharide O-antigen Biosynthetic
RI Locus of Leptospira borgpetersenii Serovar Hardjovitis."
RL Submitted (Jul-1998) to the EMBL/GenBank/DBJ databases.
CR EMBL: AF078135; AAC12950.1; -.
KW Hypothetical protein.
SQ SEQUENCE 360 AA: 41545 MW: 52201EE3 CRC32:
      67.0% Score 63; DB 2; Length 360;
      Best Local Similarity 56.7%; Pred. No. 4.07e-01;
      Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Db 129 FIKRVSVNVIHG 139
      ||||| |||||
QY 2 FIKRVSVNVIHG 13

RESULT 3 PRELIMINARY: PRT: 367 AA.
ID Q92N07
AC Q92N07
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DI 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE POLLEN MAJOR ALLERGEN 1-2.
OS Juniperus ashei (Dark white cedar).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC euphyllophytes; Spermatophyta; Coniferopsida; Coniferales;
OC Taxodiaceae; Juniperus.
RN 1
RP SEQUENCE FROM N.A.
RA MIDORO-HORTUTI T.M., GOLDBLUM R.M., KUROSKY A., WOOD T.G.,
RA BROOKS E.G.:
RI "Molecular cloning of mountain cedar (Juniperus ashei) pollen major
RT allergen, Jun a 1."
RL Submitted (Nov-1998) to the EMBL/GenBank/DBJ databases.
CR EMBL: AF106663; AAC03609.1; -.
DR EMBL: AF106662; AAC03608.1; -.
LP MENDEL: 36544; Juncas:1088;36544.
CR MENDEL: 36545; Juncas:1088;36545.
SQ SEQUENCE 367 AA: 19224 MW: 40209630 C 132;

Query Match 64.9% Score 61; DB 10; Length 367;
Best Local Similarity 44.2%; Pred. No. 4.07e-01;
Matches 6; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Db 129 FIKRVSVNVIHG 141
      ||||| |||||
QY 1 FIKRVSVNVIHG 13

RESULT 4 PRELIMINARY: PRT: 375 AA.
ID Q96385
AC Q96385
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DI 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE CHAO1 PRECURSOR.
OS Chamaecyparis obtusa.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC euphyllophytes; Spermatophyta; Coniferopsida; Coniferales;
OC Taxodiaceae; Chamaecyparis.
RN 1
RP SEQUENCE FROM N.A.
RA SUZUKI M., KOMIYAMA N., ITOH M., SONE T., KUNO K., TAKAGI I.,
RA OHTA N.:
RI "Purification, characterization and molecular cloning of Chao 1, a
RT major allergen of Chamaecyparis obtusa (Japanese cypress) pollen."
RL Mol. Immunol. 33:451-460(1996).

```

```

DR EMBL: D45404; BAAC8246.1; -.
DR MENDEL: 7625; ChaoB:1088;7626.
DR PFAM: PF00544; pec_lyase.1;
DR PRINTS: PRO0807; AMBALLERGEN.
KW Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 375 CHAO1.
SQ SEQUENCE 375 AA: 40258 MW: A0981492 CRC32:

Query Match 64.9% Score 61; DB 10; Length 375;
Best Local Similarity 53.8%; Pred. No. 4.07e-01;
Matches 7; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 129 LFMRTVSHVILHG 141
      ||||| |||||
QY 1 VFIKRVSVNVIHG 13

RESULT 5 PRELIMINARY: PRT: 390 AA.
ID Q65388
AC Q65388
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DI 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE F12F.22 PROTEIN.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC euphyllophytes; Spermatophyta; Magnoliopsida; eudicotyledons;
OC core eudicots; Rosidae; eurosids 1; Brassicales; Brassicaceae;
OC Arabidopsis.
RN 1
RP SEQUENCE FROM N.A.
RA VYSOTSKAIA V.S., OSBORNE B.L., SCHWARTZ J.R., TORIUMI M., YU S.,
RA KWAN A., GIL O., LIU S., BUEHLER E., GUNWAY A.B., GUNWAY A.R.,
RA DEWAR K., FENG J., KIM C., KURTZ D., PALM C.J., LI Y., SHINN P.,
RA SUN H., DAVIS R.W., ECKER J.R., FEDEMSPIEL N.A., THEILIGIS A.:
RL Submitted (May-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC002131; AAC17625.1; -.
DR MENDEL: 29324; Arabid:1088;29324.
DR PFAM: PF00544; pec_lyase.1;
DR PRINTS: PRO0807; AMBALLERGEN.
SQ SEQUENCE 390 AA: 43354 MW: E0322F54 CRC32:

Query Match 62.5% Score 58; DB 10; Length 390;
Best Local Similarity 72.7%; Pred. No. 1.1e-01;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Db 144 LFMRTVSHVILHG 154
      ||||| |||||
QY 3 IKRVSNNVIHG 13

RESULT 6 PRELIMINARY: PRT: 783 AA.
ID Q92H39
AC Q92H39
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DI 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE BETA-N-ACETYLGLUCOSAMINIDASE.
OS Pseudoalteromonas sp. S9.
OC Bacteria; Proteobacteria; gamma subdivision; Alteromonadaceae;
OC Pseudoalteromonas.
RN 1
RP SEQUENCE FROM N.A.
RA TECHKARNANARUK S., GOODMAN A.E.:
RI "Cloning, sequence analysis and characterisation of genes involved in
RT chitin degradation of a marine bacterium, Pseudoalteromonas sp. strain
RL S9."
RL Microbiology 130:0(1999).

```


Query Match
Best Local Similarity 50.0% Score 55 DB 10 Length 1894
Matches 5: Conservative 3: Mismatches 3: Indels 0: Gaps 0:

DB 1693 YAEVYVYVHSG 1704
QY 2 FIKRVSNVHSG 13

RESULT 12
ID Q82590
AC Q82590
DI 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DE 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
GN MAT1
OS Lepocinclis buetschlii.
OC Eukaryota; Eulenozoa; Euglenida; Euglenales; Lepocinclis.
RN [1]
RA DOEISEN N.A., THOMPSON M.D., HALLICK R.B.,
RA SEQUENCH FROM N.A.
RA "A maturase-encoding group III twintron is conserved in deeply rooted

Query Match 58.5% Score 55 DB 10 Length 1894
Best Local Similarity 50.0% Score 55 DB 10 Length 1894
Matches 5: Conservative 3: Mismatches 3: Indels 0: Gaps 0:

DB 1693 YAEVYVYVHSG 1704
QY 2 FIKRVSNVHSG 13

RESULT 12
ID Q82590
AC Q82590
DI 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DE 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
GN MAT1
OS Lepocinclis buetschlii.
OC Eukaryota; Eulenozoa; Euglenida; Euglenales; Lepocinclis.
RN [1]
RA DOEISEN N.A., THOMPSON M.D., HALLICK R.B.,
RA SEQUENCH FROM N.A.
RA "A maturase-encoding group III twintron is conserved in deeply rooted

RT Euklenoid species: Are group III introns the chicken or the egg?
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: Z99834; CAB16959.1; ...
SQ SEQUENCE 451 AA: 52368 MW: RE43A2F CRC32:
Query Match 57.4% Score 54 DB 10 Length 451
Best Local Similarity 50.0% Score 54 DB 10 Length 451
Matches 5: Conservative 4: Mismatches 1: Indels 0: Gaps 0:
DB 324 IFLPRVXN1 333
QY 1 VFIKRVSNV 10
RESULT 13
ID P72994
AC P72994
DI 01-FEB-1997 (TREMBlrel. 02, Created)
DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)
DE 01-JAN-1999 (TREMBlrel. 03, Last annotation update)
GN VAPB
OS Synechocystis sp. (strain PCC 6803)
OC Bacteria; Cyanobacteria; Chromococcales; Synechocystis
RN [1]
RA SEQUENCH FROM N.A.
RA STRAIN-PCC6803;
RA TABATA S.;
RA Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases
RN [2]
RA SEQUENCH FROM N.A.
RA STRAIN-PCC6803;
RA MEDLINE: 97061201.
RA MIYAJIMA N., HIROSAWA M., SUGIURA M., SASAKOTO S., KIMURA T.,
RA HOSOUCHI T., MATSUNO A., MURAKI A., NAKAZAKI N., NARUG K., OKUMURA S.,
RA SHIMPO S., IAKEUCHI C., WADA T., WATANABE A., YAMADA M., YASUDA M.,
RA TABATA S.;
RA "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC 6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions to
RT DNA Res. 3:109-136(1996).
DR EMBL: D9302; BAA17013.1;
SQ SEQUENCE 75 AA: 8539 MW: 9F45F6C CRC32:
Query Match 56.4% Score 53 DB 2 Length 75
Best Local Similarity 45.5% Score 53 DB 2 Length 75
Matches 5: Conservative 0: Mismatches 1: Indels 0: Gaps 0:
DB 28 VYIKKIGSTII 48
QY 1 VFIKRVSNV 11
RESULT 14
ID Q9WZIC
AC Q9WZIC
DI 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DE 01-NOV-1999 (TREMBlrel. 12, Last annotation update)
GN TM0823
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
RN [1]
RA SEQUENCH FROM N.A.
RA MEDLINE: 99287316.
RA NELSON K.E., CLAYTON R.A., GILL S.R., GWINN M.L., DODSON R.J.,
RA HAFT D.H., HICKEY E.K., PETERSON J.D., NELSON W.C., KETCHUM K.A.,
RA McDONALD L., UTTERBACK T.R., MALEK J.A., LINHER K.D., GARRETT M.M.,
RA STEWART A.M., COITON M.D., PRATT M.S., PHILLIPS C.A., RICHARDSON D.,
RA HEIDELBERG J., SUTTON G.G., FLEISCHMANN R.D., WHITE O., SALZBERG S.L.,
RA SMITH H.O., VENTER J.C., FRASER C.M.

Query Match 58.5% Score 55 DB 10 Length 1894
Best Local Similarity 50.0% Score 55 DB 10 Length 1894
Matches 5: Conservative 3: Mismatches 3: Indels 0: Gaps 0:

DB 1693 YAEVYVYVHSG 1704
QY 2 FIKRVSNVHSG 13

RESULT 12
ID Q82590
AC Q82590
DI 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DE 01-NOV-1998 (TREMBlrel. 08, Last annotation update)
GN MAT1
OS Lepocinclis buetschlii.
OC Eukaryota; Eulenozoa; Euglenida; Euglenales; Lepocinclis.
RN [1]
RA DOEISEN N.A., THOMPSON M.D., HALLICK R.B.,
RA SEQUENCH FROM N.A.
RA "A maturase-encoding group III twintron is conserved in deeply rooted

*Evidence for lateral gene transfer between Archaea and bacteria from
RT genome sequence of *Thermotoga maritima*.*

NC Nature 399:323-329(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA NELSON K.E., CLAYTON R.A., GILL S.R., GWINN M.L., DODSON R.J.,
RA HART D.H., HICKEY E.K., PETERSON J.D., NELSON W.C., KEICHUM K.A.,
RA McDONALD L., JUTERBACK T.R., MALEK J.A., LINHER K.D., GARRETT M.M.,
RA STEWART A.M., COTTON M.D., PRATT M.S., PHILLIPS C.A., RICHARDSON D.,
RA HEIDELBERG S., SUTTON G.G., FLEISCHMANN R.D., WHITE C., SALZBERG S.L.,
RA SMITH H.O., VENTER J.C., FRASER C.M.
RC Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AE001749; AAC35905.1;
SC SEQUENCE 190 AA: 22492 MW: B32D7260 CRC32:

Query Match 56.4% Score 53 DB 2 Length 317
Best Local Similarity 58.3% Pred. No. 1.79e-01
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

DB 126 FIKRVES:IEEG 137
|||||
QY 2 FIKRVSNV:IHG 13

RESULT 15
ID 066883 PRELIMINARY: PRT: 317 AA.
AC 066883:
DT 01-AUG-1998 (TrEMBLrel. 07, Created);
DI 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DE 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
RE THYMIDYLATE SYNTHASE COMPLEMENTING PROTEIN.
GN THY.
GS Aquifex aeolicus.
OC Bacteria: Aquificales: Aquificaceae: Aquifex.
RN [1]
RF SEQUENCE FROM N.A.
RC STRAIN-VFS:
FX MEDLINE: 98196566.
RA DECKER G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AJJAY M., HUBER R.,
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
aeolicus".
RC Nature 392:353-358(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-VFS:
RA DECKER G., WARREN P.V., GAASTERLAND T., YOUNG W.G., LENOX A.L.,
RA GRAHAM D.E., OVERBECK R., SNEAD M.A., KELLER M., AJJAY M., HUBER R.,
RA FELDMAN R.A., SHORT J.M., OLSON G.J., SWANSON R.V.;
PL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AE000700; AAC06547.1;
SC SEQUENCE 317 AA: 37542 MW: 6E749CF4 CRC32:

Query Match 56.4% Score 53 DB 2 Length 317
Best Local Similarity 60.0% Pred. No. 1.79e-01
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

DB 229 VFKKIENLI 238
|||||
QY 1 VFIKRVSNVI 10

Search completed: Mon Jun 19 16:16:51 2000
Job time : 11 secs.

(X)
 1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

```

MPsrch_pp      protein - protein database search, using Smith-Waterman algorithm
               Mon Jun 19 16:05:10 2000;  Maspar time 2.77 Seconds
               102.539 Million cell updates/sec
               Regular output not generated.

```

```
Title: >US-09-142-524A-6
Description: (1-12) from US09142524A.pep
Perfect Score: 72
Sequence: 1 ISLKLTSGKIAS 12
```

Scoring table:

| |
|---------|
| PAM 150 |
| Gap 15 |

searched: 18953 seqs, 23685136 residues

Post-processing: Minimum watch 08
listing first 45 summaries

Database: a - contigs35
1: genes35

Statistics: Mean 16.610; Variance 49.150; scale 0.338

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARY

| Result No. | Score | Query | | Length | DB | ID | Description | Pred. No. |
|------------|-------|-------|-----|--------|--------|----|---------------------------|-----------|
| | | Match | % | | | | | |
| 1 | 72 | 100.0 | 15 | 1 | R79799 | | Japan cedar pollen mat | 2.46e-01 |
| 2 | 72 | 100.0 | 15 | 1 | W57770 | | Residues 341-355 of Cr | 2.46e-01 |
| 3 | 72 | 100.0 | 17 | 1 | R81583 | | Cedar pollen allergen | 2.46e-01 |
| 4 | 72 | 100.0 | 17 | 1 | R90338 | | Sugi allergen protein | 2.46e-01 |
| 5 | 72 | 100.0 | 80 | 1 | W27389 | | Multi-epitope peptide | 2.46e-01 |
| 6 | 72 | 100.0 | 450 | 1 | R95971 | | Japanese cedar pollen res | 2.46e-01 |
| 7 | 72 | 100.0 | 514 | 1 | R93799 | | Japan cedar pollen Cr | 2.46e-01 |
| 8 | 72 | 100.0 | 514 | 1 | R74333 | | Japanese cedar pollen | 2.46e-01 |
| 9 | 72 | 100.0 | 514 | 1 | R69792 | | Japanese cedar pollen | 2.46e-01 |
| 10 | 72 | 100.0 | 514 | 1 | R33690 | | Japanese cedar pollen | 2.46e-01 |
| 11 | 72 | 100.0 | 514 | 1 | R81586 | | Cedar pollen allergen | 2.46e-01 |
| 12 | 65 | 90.3 | 17 | 1 | R81577 | | Cedar pollen allergen | 1.67e-00 |
| 13 | 65 | 90.3 | 17 | 1 | R41594 | | Cedar pollen allergen | 1.67e-00 |
| 14 | 61 | 84.7 | 12 | 1 | W4301 | | Japanese cedar pollen | 4.93e-00 |
| 15 | 61 | 84.7 | 15 | 1 | R97940 | | Japanese cedar pollen mat | 4.93e-00 |
| 16 | 59 | 81.9 | 20 | 1 | W42197 | | T-cell epitope peptide | 8.12e-00 |
| 17 | 59 | 81.9 | 514 | 1 | W04346 | | Chamaecyparis obtusa p | 8.12e-00 |
| 18 | 59 | 81.9 | 514 | 1 | W42122 | | Japanese cypress polle | 8.12e-00 |
| 19 | 54 | 75.0 | 35 | 1 | W83342 | | Sugi allergen protein | 3.37e-01 |
| 20 | 54 | 75.0 | 35 | 1 | W83344 | | Sugi allergen protein | 3.37e-01 |
| 21 | 53 | 73.6 | 11 | 1 | W4302 | | Japanese cedar pollen | 3.97e-01 |
| 22 | 53 | 73.6 | 12 | 1 | W4303 | | Japanese cedar pollen | 3.97e-01 |
| 23 | 53 | 73.6 | 33 | 1 | W83343 | | Sugi allergen protein | 3.97e-01 |

ALIGNMENTS

RESULT 1
ID R97935 standard; peptide; 15 AA.
AC R97939;
AD R97939;
DE 16-AUG-1996 (first entry)
DE Japan cedar pollen mature allergen Cry 3 II amino acids 341-355.
DE Allergen; epitope: overlapping peptide; Cry 3 II; cedar pollen;
KW Sugii pollinosis; diagnosis: treatment.
KW Cryptomeria japonica.
PN J08047392-A.
PD J08047392-A.
PD 20-FEB-1996.
PF 07-NOV-1994: 297840.
PF 05-NOV-1993: Cp-276773.
PR 26-MAY-1994: JP-134858.
PA (MEIP) MELJI MILK PROD CO. I.
DR WPI: 96-166245/17.
PT Japan cedar pollen allergen (Cry 3 II) epitope - comprises at least:
PT Part of specified 450 amino acid protein
PS Claim 8; Fig 5; 17pp: Japanese.
PS R57871-R57960 are overlapping peptides used for the epitope mapping
CC of the Japan cedar pollen allergen Cry 3 II. Cry 3 II and allergenic
CC peptides of it are useful in the diagnosis, prevention and treatment
CC of Sugii pollinosis, the allergic reaction to Japan cedar pollen.
CC Significant regions of the allergen were identified using the
CC overlapping peptides of the full epitope derived from a Cry 3 II:
CC antigen-specific T cell line. Amino acids 66-80 (R97884) and 186-290
CC (R97890) of the full mature 460 amino acid allergen are the most
CC allergenic of the 90 peptides tested.
S0 Sequence 15 AA:

```
Query Match      100.0%: Score 72: DB 1: Length 15:
Best Local Similarity 100.0%: Pred. No. 2.45e-01:
Matches 12: Conservative 0: Mismatches 0: Indels
```

```
Db 4 ISKLTSGKIAS 15
    1111111111111111
Ov 1 ISKLTSGKIAS 12
```

| | | |
|--------|--------------------------------|----------------------------------|
| RESULT | 2 | W57770 standard: peptide: 15 AA. |
| AC | W57770: | |
| CD | 17-SEP-1998 | (first entry) |
| DE | Residues 341-355 of Cry j 2 | |
| KW | Cry j 2: Japanese cedar pollen | |
| KW | H:A class, 11 molecule. | |
| PN | Cryptomeria japonica. | |
| PS | W5987C902-A. | |

12-MAY-1998.
 DE 12-NOV-1997: 304129.
 DE 13-NOV-1996: JP-302053.
 PA (MEIP) MEIJI MILK PROD CO LTD.
 PI Dairiki K, Kino K, Kume A, Sone T.
 DR WPI: 98-397677/26.
 PT Peptides derived from Japanese cedar pollen antigens are
 PI immunotherapeutic agents - useful for allergy treatment and typing
 PI HLA class II molecules in allergy sufferers
 PS Claim 5: Page 3: 50pp: Japanese.
 CC This sequence represents residues 341-355 of the Cry j 2 protein, and
 CC is a peptide of the invention. The peptides are derived from Japanese
 CC cedar pollen antigens, and are used as immunotherapeutic agents in the
 CC treatment of allergy. The peptides can be used for identification and
 CC typing of the particular HLA class II molecules in an allergy sufferer,
 CC and also for peptide immunotherapy of an allergy. Using these peptides,
 CC the immunotherapy can be targeted more specifically to the requirements
 CC of the individual patient, allowing more effective treatment of an
 CC allergy, including those patients for whom treatment with a conventional
 CC immunotherapeutic agent is ineffective.
 SQ Sequence: 15 AA:
 Query Match 100.0% Score 72: DB 1: Length 15:
 Best Local Similarity 100.0% Pred. No. 2,46e-01:
 Matches 12: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 DB 4 ISLKLTSKGKIAS 15
 QY 1 ISLKLTSKGKIAS 12
 RESULT 3
 ID R81583 standard: Peptide: 17 AA.
 AC R81583:
 CI 24-MAY-1996 (first entry)
 DE Cedar pollen allergen peptide 11 (T-cell epitope).
 KW Cedar, Cryptocarya japonica; pollen; allergen; immunoglobulin E;
 KW IgE; T-cell epitope; antibody; pollinosis therapy; immunotherapy.
 OS Synthetic.
 PN EP-700929-A2.
 PD 13-MAR-1996.
 PF 08-SEP-1995: 306295.
 PR 10-SEP-1994: JP-242137.
 PR 14-JUN-1993: JP-280221.
 PR 14-JUN-1993: JP-280204.
 PA (HAYB) HAYASHIBARA SEIBUSU KAGAKU.
 PI Hino K, Saito S, Taniuchi Y.
 DR WPI: 96-140976/45.
 PI New peptide(s) derived from cedar pollen allergens - activate
 PI allergen-specific T-cells, but not allergen-specific IgE antibodies.
 PI used for treating cedar pollinosis
 PS Claim 4: Page 29: 36pp: English.
 CC Synthetic peptides based on portions of cedar pollen allergens A
 CC (R81586) and B (R81587) were tested for their ability to activate
 CC cedar allergen-specific T-cells, but not allergen-specific IgE
 CC antibodies. 6 peptides (R81580-R81585) were identified as T-cell
 CC epitopes. These peptides, plus subsequences (R81573-79) essential
 CC for T-cell recognition, and homologous peptides (R81588-96) can
 CC be used as immunotherapeutic agents to treat or prevent cedar
 CC pollinosis, avoiding side-effects such as anaphylaxis.
 SQ Sequence: 17 AA:
 Query Match 100.0% Score 72: DB 1: Length 17:
 Best Local Similarity 100.0% Pred. No. 2,46e-01:
 Matches 12: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 DB 2 ISLKLTSKGKIAS 13
 QY 1 ISLKLTSKGKIAS 12

W80348:
 AC 11-JAN-1999 (first entry)
 DE Sugi allergen protein Cryj2 derived epitope for T cells.
 DE T cell epitope; sugi allergen proteins (Cryj1; Cryj2); treatment;
 KW sugi-pollinosis; allergic reaction; pollen.
 OS Synthetic.
 PN J10259198-A.
 PD 29-SEP-1998.
 PF 22-DEC-1997: 353448.
 PR 24-DEC-1996: JP-343441.
 PA (HAYB) HAYASHIBARA SEIBUSU KAGAKU.
 PI (SANY) SANKYO CO LTD.
 DR WPI: 98-577037/49.
 PT A linked T cell epitope pepi used for the treatment of
 PI sugi-pollinosis
 PS Claim 7: Page 18: 21pp: Japan.
 CC W80339-58 represent epitopes T cells derived from the sugi allergen
 CC proteins Cryj1 (W80339-44, W80340-53 and W80356-58) and Cryj2 (W80345-49
 CC and W80354-55). The peptides are useful for the treatment of
 CC sugi-pollinosis, an allergic reaction of the body to pollen.
 SQ Sequence: 17 AA:
 Query Match 100.0% Score 72: DB 1: Length 17:
 Best Local Similarity 100.0% Pred. No. 2,46e-01:
 Matches 12: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 DB 1 ISLKLTSKGKIAS 12
 QY 1 ISLKLTSKGKIAS 12
 RESULT 5
 ID W27369 standard: peptide: 80 AA
 AC W27369:
 CI 24-MAR-1998 (first entry)
 DE Multi-epitope peptide used as immunotherapeutic agent #1.
 KW Multi-epitope peptide; immunotherapeutic agent; allergic disease.
 KW T-cell epitope region; allergen; lymphocyte; immunoglobulin E.
 OS Synthetic.
 PN WO9732600-A1.
 PD 12-SEP-1997.
 PF 10-MAR-1997: 300740.
 PR 10-MAR-1996: JP-080702.
 PA (MEIP) MEIJI MILK PROD CO LTD.
 PI Dairiki K, Iwama A, Kino K, Kume A, Sone T.
 DR WPI: 97-479495/43.
 PI Peptide immunotherapeutic agent to treat allergic diseases
 PI contains multi-epitope peptide containing T cell epitope regions
 PI from different allergens
 PS Claim 6: Page 31: 58pp: Japanese.
 CC The present sequence represents a multi-epitope peptide which is used as
 CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2
 CC or more different allergens (preferably linked via arginine or lysine
 CC dimers), where the T cell epitope regions have a positivity index
 CC greater than 100 as measured in a patient group responding to the
 CC allergen; have at least 70% reactivity with lymphocytes from patients
 CC responding to the allergen; and are not reactive with immunoglobulin E
 CC (IgE) antibodies from patients responsive to the allergen. The agent can
 CC be used to prevent and treat a wide variety of allergic diseases, e.g. by
 CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
 SQ Sequence: 80 AA:
 Query Match 100.0% Score 72: DB 1: Length 80:
 Best Local Similarity 100.0% Pred. No. 2,46e-01:
 Matches 12: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
 DB 52 ISLKLTSKGKIAS 53
 QY 1 ISLKLTSKGKIAS 12
 RESULT 6
 ID R6379 standard: Protein: 460 AA.

Db 398 ISKLTSCKIAS 409
|||||

QY 1 ISKLTSCKIAS 12

RESULT 12
ID R81577 standard: Peptide: 11 AA.

AC R81577 (first entry)

DE Cedar pollen allergen peptide 5

KW Cedar; Cryptomeria japonica; pollen; allergen; immunoglobulin E

OS IgE antibody; pollinosis; therapy; immunotherapy

PN Synthetic

PD EP-70029-A2

PF 13-MAR-1996: 306295

PR 08-SEP-1995: JP-242137

PR 10-SEP-1994: JP-242137

PR 14-JUL-1995: JP-200221

PR 14-JUL-1995: JP-200204

PA (HAYB) HAYASHIBARA SEIBUTSU KAKAKU

PI Hino K, Saito S, Taniuchi Y

DR WPI: 96-140976/15

PT New peptide(s) derived from cedar pollen allergens - activate

PI allergen-specific T-cells, but not allergen-specific IgE antibodies

PS Claim 2: Fig 4: 9pp: English

CC The sequence is of a Japanese cedar pollen allergen Cry j

CC in the protein and its fragments can be used for diagnosis and

CC treatment of Japanese cedar pollinosis and to identify similar

CC sequences in other plants

CC See also R81586-95

CC Sequence 514 AA

QY 1 ISKLTSCKIAS 12

Query Match 100.0% Score 72: DB 1: Length 514

Best Local Similarity 100.0% Pred. No. 2.46e-01

Matches 12: Conservative 0: Mismatches 0: Indels 0: Gaps 0

Db 398 ISKLTSCKIAS 409

|||||

QY 1 ISKLTSCKIAS 12

RESULT 11

ID R81586 standard: Protein: 514 AA

AC R81586 (first entry)

DE Cedar pollen allergen A

KW Cedar; pollen; allergen; immunoglobulin E; IgE; T-cell epitope;

OS antibody; pollinosis; therapy; immunotherapy

PN Cryptomeria japonica

PD EP-70029-A2

PF 13-MAR-1996

PR 08-SEP-1995: 306295

PR 10-SEP-1994: JP-242137

PR 14-JUL-1995: JP-200221

PR 14-JUL-1995: JP-200204

PA (HAYB) HAYASHIBARA SEIBUTSU KAKAKU

PI Hino K, Saito S, Taniuchi Y

DR WPI: 96-140976/15

PT New peptide(s) derived from cedar pollen allergens - activate

PI allergen-specific T-cells, but not allergen-specific IgE antibodies

PS Claim 5: Page 29-30: 36pp: English

CC Synthetic peptides based on portions of cedar pollen allergens A

CC (R81586) and B (R81587) were tested for their ability to activate

CC cedar allergen-specific T-cells, but not allergen-specific IgE

CC antibodies. 6 peptides (R81580-R81585) were identified as T-cell

CC epitopes. These peptides, plus subsequences (R81573-79) essential

CC for T-cell recognition, and homologous peptides (R81588-95) can

CC be used as immunotherapeutic agents to treat or prevent cedar

CC pollinosis, avoiding side-effects such as anaphylaxis

CC Sequence 514 AA

Db 398 ISKLTSCKIAS 409

|||||

QY 1 ISKLTSCKIAS 12

RESULT 10

ID R81577 standard: Protein: 514 AA

AC R81577 (first entry)

DE Cedar pollen allergen peptide 5

KW Cedar; Cryptomeria japonica

OS IgE antibody; pollinosis; therapy; immunotherapy

PN Synthetic

PD EP-70029-A2

PF 13-MAR-1996: 306295

PR 08-SEP-1995: JP-242137

PR 10-SEP-1994: JP-242137

PR 14-JUL-1995: JP-200221

PR 14-JUL-1995: JP-200204

PA (HAYB) HAYASHIBARA SEIBUTSU KAKAKU

PI Hino K, Saito S, Taniuchi Y

DR WPI: 96-140976/15

PT New peptide(s) derived from cedar pollen allergens - activate

PI allergen-specific T-cells, but not allergen-specific IgE antibodies

PS Claim 2: Fig 4: 9pp: English

CC The sequence is of a Japanese cedar pollen allergen Cry j

CC in the protein and its fragments can be used for diagnosis and

CC treatment of Japanese cedar pollinosis and to identify similar

CC sequences in other plants

CC See also R81586-95

CC Sequence 514 AA

QY 1 ISKLTSCKIAS 12

Query Match 100.0% Score 72: DB 1: Length 514

Best Local Similarity 100.0% Pred. No. 2.46e-01

Matches 12: Conservative 0: Mismatches 0: Indels 0: Gaps 0

Db 398 ISKLTSCKIAS 409

|||||

QY 1 ISKLTSCKIAS 12

RESULT 11

ID R81586 standard: Protein: 514 AA

AC R81586 (first entry)

DE Cedar pollen allergen A

KW Cedar; pollen; allergen; immunoglobulin E; IgE; T-cell epitope;

OS antibody; pollinosis; therapy; immunotherapy

PN Cryptomeria japonica

PD EP-70029-A2

PF 13-MAR-1996

PR 08-SEP-1995: 306295

PR 10-SEP-1994: JP-242137

PR 14-JUL-1995: JP-200221

PR 14-JUL-1995: JP-200204

PA (HAYB) HAYASHIBARA SEIBUTSU KAKAKU

PI Hino K, Saito S, Taniuchi Y

DR WPI: 96-140976/15

PT New peptide(s) derived from cedar pollen allergens - activate

PI allergen-specific T-cells, but not allergen-specific IgE antibodies

PS Claim 5: Page 29-30: 36pp: English

CC Synthetic peptides based on portions of cedar pollen allergens A

CC (R81586) and B (R81587) were tested for their ability to activate

CC cedar allergen-specific T-cells, but not allergen-specific IgE

CC antibodies. 6 peptides (R81580-R81585) were identified as T-cell

CC epitopes. These peptides, plus subsequences (R81573-79) essential

CC for T-cell recognition, and homologous peptides (R81588-95) can

CC be used as immunotherapeutic agents to treat or prevent cedar

CC that contain 1 or more amino acid substitution(s) can be used
 CC as immunotherapeutic agents to treat or prevent cedar pollinosis,
 CC avoiding side-effects such as anaphylaxis.

SQ Sequence 17 AA:

Query Match 90.3%; Score 65; DB 1; Length 17;
 Best Local Similarity 100.0%; Pred. No. 1.97e+00;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 3 SLKLTSGKIAS 13

IIIIIIIIII

QY 2 SLKLTSGKIAS 12

RESULT 14

ID W14301 standard; peptide: 12 AA.

AC W14301

DE 30-APR-1997 (first entry)

DE Japanese cedar pollen allergen-derived peptide 20.

KW Japanese cedar; sugi; pollen; pollinosis; pollenosis; allergen;

KW allergen; T-cell epitope; TCR; T-cell receptor; activation;

KW Immune tolerance; Cry1.

CS Cryptomeria japonica.

CS Synthetic.

PN J08334391-A.

PE 17-DEC-1996.

PF 18-JUL-1995; 181438.

FF 07-APR-1995; JP-322519.

FA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU.

FA (SANY) SANKYO CO LTD.

DE WPI: 97-095487/09.

P1 Peptide allergen derived from Japanese cedar pollen - causes T cell

P1 response specific to cedar pollen; for treatment of pollinosis

PS Claim 29; Page 20; 21pp; Japanese.

CC The present sequence is one of 24 claimed peptides which were

CC synthesised based on Japanese cedar pollen sequences. This peptide

CC was shown to have Japanese cedar pollen antigen T cell epitope

CC activity by using T cells isolated from a cedar pollinosis patient.

CC The peptide produces little or no anaphylaxis.

SQ Sequence 12 AA:

Query Match 94.7%; Score 61; DB 1; Length 12;
 Best Local Similarity 100.0%; Pred. No. 4.90e+00;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 LKLTSGKIAS 10

IIIIIIII

QY 3 LKLTSGKIAS 12

RESULT 15

ID R97940 standard; peptide: 15 AA.

AC R97940

DE 16-AUG-1996 (first entry)

DE Japan cedar pollen mature allergen Cry j II amino acids 346-350.

KW Allergen; epitope; overlapping peptide; Cry j II; cedar pollen;

KW Sugi pollinosis; diagnosis; treatment.

CS Cryptomeria japonica.

PN J08347392-A.

PE 20-FEB-1996.

PF 07-NOV-1994; 297840.

FF 05-NOV-1993; JP-276773.

PR 25-MAY-1994; JP-134858.

PA (MEIP) MELIJI MILK PROD CO LTD.

DE WPI: 96-166249/17.

P1 Japan cedar pollen allergen Cry j II epitope - comprises at least

P1 part of specified 460 amino acid protein

PS Claim 8; Fig 5; 17pp; Japanese.

CC R97871-R97960 are overlapping peptides used for the epitope mapping

CC of the Japan cedar pollen allergen Cry j II. Cry j II and allergenic

CC peptides of it are useful in the diagnosis, prevention and treatment

CC of Sugi pollinosis, the allergic reaction to Japan cedar pollen.

CC Significant regions of the allergen were identified using the

CC overlapping peptides of the full epitope derived from a Cry j II
 CC antigen-specific T cell line. Amino acids 65-80 (R97884) and 185-200
 CC (R978908) of the full mature 460 amino acid allergen are the most
 CC allergenic of the 90 peptides tested.

SQ Sequence 15 AA:

Query Match 84.7%; Score 61; DB 1; Length 15;
 Best Local Similarity 100.0%; Pred. No. 4.90e+00;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 LKLTSGKIAS 10

IIIIIIIIII

QY 3 LKLTSGKIAS 12

Search completed: Mon Jun 19 16:05:21 2000

Job time : 11 secs.

[REDACTED]

(TM)

Release 3.1A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd
Mfsrcch_pp protein - protein database search, using Smith-Waterman algorithm
Run on: Tue Jun 20 13:34:47 2000; MasPar time 2.23 seconds
Tabular output not generated. 77.802 Million cell updates/sec

Title: >US-09-142-524A-6
Description: (1-12) from US09142524A.pep
Perfect Score: 72
Sequence: 1 ISLKLISGKIAS 12
Scoring table: PWM 150
Gap 15
Searched: 145341 seqs, 14437480 residues
Post-processing: Minimum Match 0%
Listing first 45 summaries
Database: a-issued
1:5A_COMB 2:5B_COMB 3:5_C01 4:PCT_COMB 5:backfiles:
Statistics: Mean 15.504; Variance 46.068; scale 0.337

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARY

| Result No. | Score | Query Match | Length | ID | Description | Pred. No. |
|------------|-------|-------------|--------|----|----------------------------------|-----------|
| 1 | 47 | 65.3 | 572 | 2 | US-08-993-Sequence 7, Applicatio | 8.72e+01 |
| 2 | 47 | 65.3 | 572 | 2 | US-09-032-Sequence 7, Applicatio | 8.72e+01 |
| 3 | 46 | 63.9 | 263 | 1 | US-08-407-Sequence 2, Applicatio | 1.12e+02 |
| 4 | 46 | 63.9 | 284 | 2 | US-08-719-Sequence 8, Applicatio | 1.12e+02 |
| 5 | 46 | 63.9 | 285 | 2 | US-08-719-Sequence 2, Applicatio | 1.12e+02 |
| 6 | 45 | 63.9 | 285 | 2 | US-08-719-Sequence 6, Applicatio | 1.12e+02 |
| 7 | 46 | 63.9 | 286 | 1 | US-08-346-Sequence 16, Applicati | 1.12e+02 |
| 8 | 46 | 63.9 | 286 | 2 | US-08-719-Sequence 4, Applicatio | 1.12e+02 |
| 9 | 46 | 63.9 | 286 | 1 | US-07-721-Sequence 2, Applicatio | 1.12e+02 |
| 10 | 46 | 63.9 | 286 | 4 | PCT-US91-0Sequence 2, Applicatio | 1.12e+02 |
| 11 | 46 | 63.9 | 286 | 1 | US-08-339-Sequence 2, Applicatio | 1.12e+02 |
| 12 | 46 | 63.9 | 359 | 1 | US-08-503-Sequence 2, Applicatio | 1.12e+02 |
| 13 | 46 | 63.9 | 359 | 2 | US-08-576-Sequence 2, Applicatio | 1.12e+02 |
| 14 | 46 | 63.9 | 359 | 3 | US-08-899-Sequence 2, Applicatio | 1.12e+02 |
| 15 | 46 | 63.9 | 359 | 2 | US-08-972-Sequence 2, Applicatio | 1.12e+02 |
| 16 | 46 | 63.9 | 1186 | 4 | PCT-US93-1Sequence 6, Applicatio | 1.12e+02 |
| 17 | 46 | 63.9 | 1186 | 1 | US-08-357-Sequence 6, Applicatio | 1.12e+02 |
| 18 | 46 | 63.9 | 1186 | 2 | US-08-590-Sequence 4, Applicatio | 1.12e+02 |
| 19 | 46 | 63.9 | 1186 | 1 | US-08-485-Sequence 4, Applicatio | 1.12e+02 |
| 20 | 46 | 63.9 | 1186 | 2 | US-09-184-Sequence 4, Applicatio | 1.12e+02 |
| 21 | 46 | 63.9 | 1308 | 2 | US-08-996-Sequence 2, Applicatio | 1.12e+02 |
| 22 | 45 | 62.5 | 331 | 4 | PCT-US95-1Sequence 4, Applicatio | 1.44e+02 |
| 23 | 45 | 62.5 | 331 | 1 | US-08-208-Sequence 8, Applicatio | 1.44e+02 |

| | | | | | | |
|----|----|------|------|---|----------------------------------|----------|
| 24 | 45 | 62.5 | 331 | 1 | US-8-330-Sequence 4, Applicatio | 1.44e+02 |
| 25 | 45 | 62.5 | 644 | 2 | US-8-866-Sequence 2, Applicatio | 1.44e+02 |
| 26 | 44 | 61.1 | 658 | 1 | US-08-409-Sequence 5, Applicatio | 1.85e+02 |
| 27 | 44 | 61.1 | 1098 | 1 | US-08-409-Sequence 2, Applicatio | 1.85e+02 |
| 28 | 43 | 59.7 | 1489 | 5 | 5183745-2atent No. 5183745 | 2.37e+02 |
| 29 | 43 | 59.7 | 1794 | 5 | 5183745-6atent No. 5183745 | 2.37e+02 |
| 30 | 42 | 58.3 | 179 | 2 | US-08-933-Sequence 34, Applicat | 3.03e+02 |
| 31 | 42 | 58.3 | 247 | 5 | 5455030-9atent No. 5455030 | 3.03e+02 |
| 32 | 42 | 58.3 | 318 | 2 | US-08-278-Sequence 3, Applicatio | 3.03e+02 |
| 33 | 42 | 58.3 | 401 | 2 | US-08-278-Sequence 2, Applicatio | 3.03e+02 |
| 34 | 42 | 58.3 | 460 | 1 | US-8-833-Sequence 50, Applicati | 3.03e+02 |
| 35 | 42 | 58.3 | 460 | 1 | US-8-306-Sequence 50, Applicati | 3.03e+02 |
| 36 | 42 | 58.3 | 460 | 1 | US-8-476-Sequence 50, Applicati | 3.03e+02 |
| 37 | 42 | 58.3 | 863 | 2 | US-8-566-Sequence 2, Applicatio | 3.03e+02 |
| 38 | 42 | 58.3 | 2973 | 2 | US-8-603-Sequence 7, Applicatio | 3.03e+02 |
| 39 | 42 | 58.3 | 2973 | 2 | US-8-821-Sequence 7, Applicatio | 3.03e+02 |
| 40 | 42 | 58.3 | 3075 | 2 | US-8-125-Sequence 5, Applicatio | 3.03e+02 |
| 41 | 42 | 58.3 | 3075 | 2 | US-8-460-Sequence 5, Applicatio | 3.03e+02 |
| 42 | 42 | 58.3 | 3177 | 2 | US-08-477-Sequence 4, Applicatio | 3.03e+02 |
| 43 | 41 | 56.9 | 144 | 1 | US-08-186-Sequence 10, Applicati | 3.85e+02 |
| 44 | 41 | 56.9 | 476 | 2 | US-08-850-Sequence 2, Applicatio | 3.85e+02 |
| 45 | 41 | 56.9 | 1456 | 1 | US-08-803-Sequence 2, Applicatio | 3.85e+02 |

ALIGNMENTS

RESULT 1
ID US-08-993-318A-7 STANDARD: PRT: 572 AA.
XX
AC xxxxxx
XX
D-

Sequence 7, Application US/08993318A

Sequence 7, Application US/08993318A

Patent No. 5998353

GENERAL INFORMATION:

APPLICANT: Pedersen, Anders

APPLICANT: Svendsen, Allan

APPLICANT: Schneider, Palie

APPLICANT: Rasmussen, G. the

APPLICANT: Cherry, Joel

TITLE OF INVENTION: LAC USE MUTANTS

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: NO. 5998353 No. 5998353disk of No. 5998353, America

STREET: 405 Lexington Avenue

CITY: New York

COUNTRY: USA

ZIP: 10174

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30 (EPC)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/993,318A

FILING DATE: December 18, 1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Gregq, Valeta

REGISTRATION NUMBER: 33,728

REFERENCE/DOCKET NUMBER: 5032,200-US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-867-012

TELEFAX: 212-878-9655

TELEX:

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 572 amino acids

TYPE: amino acid

STRANDEDNESS: single

| | | |
|--------|--|--|
| CC | TOPOLOGY: | linear |
| CC | MOLECULE TYPE: | protein |
| CC | SEQUENCE | 572 AA: 63747 MW: 1718990 CN: |
| SS | | |
| DE | Query Match: | 65.3%; Score 47; DB 2; Length 572; |
| XX | Best Local Similarity | 54.5%; Pred. No. 8.72e-01; |
| XX | Matches | 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0; |
| DL | 25 NLKISNGKIAP 35 | |
| QY | | |
| QY | 2 SLKLTSGKIAS 12 | |
| RESULT | 2 | |
| ID | US-09-032-315-7 | STANDARD: PRT: 572 AA. |
| XX | xxxxxx | |
| DE | Sequence 7, Application US/09032315 | |
| CC | Patent No. 545818 | |
| CC | GENERAL INFORMATION: | |
| CC | APPLICANT: Svendsen, Allan | |
| CC | APPLICANT: Xu, Feng | |
| CC | TITLE OF INVENTION: LACCASE MUTANTS | |
| CC | NUMBER OF SEQUENCES: 10 | |
| CC | CORRESPONDENCE ADDRESS: | |
| CC | ADDRESSER: NO. 5985818g NO. 5985818disk of NO. 5985818th America, Inc. | |
| CC | STREET: 405 Lexington Avenue | |
| CC | CITY: New York | |
| CC | STATE: NY | |
| CC | COUNTRY: USA | |
| CC | ZIP: 10174 | |
| CC | COMPUTER READABLE FORM: | |
| CC | MEDIUM TYPE: Diskette | |
| CC | COMPUTER: IBM Compatible | |
| CC | OPERATING SYSTEM: DOS | |
| CC | SOFTWARE: FASTSEQ for Windows Version 2.0 | |
| CC | CURRENT APPLICATION DATA: | |
| CC | APPLICATION NUMBER: US/09/032.315 | |
| CC | FILING DATE: 27-FEB-1998 | |
| CC | CLASSIFICATION: | |
| CC | ATTORNEY/AGENT INFORMATION: | |
| CC | NAME: Foxok, Carol | |
| CC | REGISTRATION NUMBER: 36,993 | |
| CC | REFERENCE/DOCKET NUMBER: 5300-200-US | |
| CC | TELECOMMUNICATION INFORMATION: | |
| CC | TELEPHONE: 212-867-0123 | |
| CC | TELEFAX: 212-878-9655 | |
| CC | INFORMATION FOR SEQ ID NO: 7: | |
| CC | SEQUENCE CHARACTERISTICS: | |
| CC | LENGTH: 572 amino acids | |
| CC | TYPE: amino acid | |
| CC | STRANDEDNESS: single | |
| CC | TOPOLOGY: linear | |
| CC | MOLECULE TYPE: protein | |
| SS | SEQUENCE 572 AA: 63747 MW: 1718990 CN: | |
| DE | Query Match: | 65.3%; Score 47; DB 2; Length 572; |
| XX | Best Local Similarity | 54.5%; Pred. No. 8.72e-01; |
| XX | Matches | 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0; |
| DL | 25 NLKISNGKIAP 35 | |
| QY | | |
| QY | 2 SLKLTSGKIAS 12 | |
| RESULT | 3 | |
| ID | US-08-407-544-2 | STANDARD: PRT: 263 AA. |
| XX | xxxxxx | |
| DE | Sequence 8, Application US/08719697 | |
| CC | Patent No. 5928888 | |
| CC | GENERAL INFORMATION: | |
| CC | APPLICANT: Whitney, Michael A. | |
| CC | TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SENSITIVE | |
| CC | TITLE OF INVENTION: AND RAPID, FUNCTIONAL IDENTIFICATION OF GENOMIC | |
| CC | TITLE OF INVENTION: POLYNUCLEOTIDES AND SECONDARY SCREENING CAPABILITIES | |
| CC | NUMBER OF SEQUENCES: 10 | |
| CC | CORRESPONDENCE ADDRESS: | |

Matches 7: Conservative 0: Mismatches 3: Indels 0: Gaps 0:

Db 24 IELDLSGK: 33
1 ISLKTSGK: 10

RESULT 7
ID US-08-346-333-16 STANDARD: PRT: 286 AA.
XX XXXXX
XX XXXXX

Sequence 16, Application US/08346333
Patent No. 597153
GENERAL INFORMATION:
APPLICANT: Botstein, David
TITLE OF INVENTION: Methods for modifying DNA and for
TITLE OF INVENTION: detecting effects of such modification on interaction of
TITLE OF INVENTION: encoded modified polypeptides with target substrates.
NUMBER OF SEQUENCES: 86
CORRESPONDENCE ADDRESS:
ADDRESS: Richard F. Treccartin
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/346,333
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/039,501
FILING DATE:
APPLICATION NUMBER: US/07/602,156
FILING DATE: 22-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Treccartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A 53469/RP
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3245
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 286 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE 286 AA: 31456 MW: 397468 CN:

Query Match: 63.9% Score 46: DB 1: Length 286:
Best Local Similarity 70.0%: Pct. Ident. 0:
Matches 7: Conservative 0: Mismatches 3: Indels 0: Gaps 0:

Db 45 IELDLSGK: 54
1 ISLKTSGK: 10

RESULT 8
ID US-08-719-697-4 STANDARD: PRT: 286 AA.
XX XXXXX

xxxxxx

Sequence 4, Application US/08719697
Sequence 4, Application US/08719697
Patent No. 592888
GENERAL INFORMATION:
APPLICANT: Whitney, Michael A.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SENSITIVE
TITLE OF INVENTION: AND RAPID, FUNCTIONAL IDENTIFICATION OF GENOMIC
TITLE OF INVENTION: POLYNUCLEOTIDES AND SECONDARY SCREENING CAPABILITIES
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESS: Fish & Richardson P.C.
STREET: 4225 Executive Square, Suite 1400
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/719,697
FILING DATE: 26-SEP-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Haile, Lisa A.
REGISTRATION NUMBER: 38,147
REFERENCE/DOCKET NUMBER: 04366/00001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619/678-5076
TELEFAX: 619/678-5099
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 286 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE 286 AA: 31570 MW: 446420 CN:

Query Match: 63.9% Score 46: DB 1: Length 286:
Best Local Similarity 70.0%: Pct. Ident. 0:
Matches 7: Conservative 0: Mismatches 3: Indels 0: Gaps 0:

Db 45 IELDLSGK: 54
1 ISLKTSGK: 10

RESULT 9
ID US-07-721-775A-2 STANDARD: PRT: 286 AA.
XX XXXXX
XX XXXXX
XX XXXXX

Sequence 2, Application US/07721775A
Sequence 2, Application US/07721775A
Patent No. 5180666
GENERAL INFORMATION:
APPLICANT: States, J. Christopher
APPLICANT: Hines, Ronald N.
TITLE OF INVENTION: NO. 5180666ak, Raymond F.
TITLE OF INVENTION: METHOD AND CELL LINE FOR TESTING
TITLE OF INVENTION: MUTAGENICITY OF A CHEMICAL
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:

SO SEQUENCE 286 AA: 31557 MW: 391925 CN:

Query Match 63.9% Score 46: DB 1: Length 286;
Best Local Similarity 70.0% Pred. No. 1:12e-02;
Matches 7: Conservative 0: Mismatches 3: Indels 0: Gaps 0:

DB 45 IEDLNSGK: 54
QY 1 SLKLTSGKI 10

RESULT 12
ID US-08-503-133A-2 STANDARD: PRT: 359 AA:

XX AC xxxxxx

DI

DE

XX

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

US-08-576-775A-2 STANDARD: PRT: 359 AA:

XX AC xxxxxx

DI

DE

XX

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

Sequence 2, Application US/08576775A

Sequence 2, Application US/08576775A

Patent No. 5849504

GENERAL INFORMATION:

APPLICANT: Gerardy-Schahn, Rita; Fukuda, Minoru;

APPLICANT: Nakayama, Jun; Eckhardt, Matthias

TITLE OF INVENTION: Isolated Polysialyl Transferases.

TITLE OF INVENTION: Nucleic Acid Molecules Coding Therefor.

TITLE OF INVENTION: Methods of

PRODUCTION AND USE

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Felie & Lynch

STREET: 805 Third Avenue

CITY: New York City

STATE: New York

COUNTRY: USA

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage

COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/576.775A

FILING DATE: 21-December-1995

CLASSIFICATION: 514

Prior APPLICATION DATA:

APPLICATION NUMBER: 08/ 3,133

FILING DATE: 17-July-1994

Prior APPLICATION DATA:

APPLICATION NUMBER: PC 94/04289

FILING DATE: 22-December-1994

ATTORNEY/AGENT INFORMATION:

NAME: Hanson, No. 584726man D.

REGISTRATION NUMBER: 0,946

REFERENCE/DOCKET NUMBER: BOER 1,500

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 688-9200

TELEFAX: (212) 688-9200

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 359 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE 359 AA: 41227 MW: 690704 CN:

Query Match 63.9% Score 46: DB 2: Length 359;

Best Local Similarity 75.0% Pred. No. 1:12e-02;

Matches 6: Conservative 2: Mismatches 0: Indels 0: Gaps 0:

DB 348 ALKLTGK 355

QY 2 SLKLTSGK 9

RESULT 14

ID US-08-899-545-2 STANDARD: PRT: 359 AA:

XX AC xxxxxx

DI

DE

XX

XX

Sequence 2, Application US/08899545

Sequence 2, Application US/08899545

Patent No. 5849504

GENERAL INFORMATION:

APPLICANT: Gerardy-Schahn, Rita; Fukuda, Minoru;

APPLICANT: Nakayama, Jun; Eckhardt, Matthias

TITLE OF INVENTION: Isolated Polysialyl Transferases.

TITLE OF INVENTION: Nucleic Acid Molecules Coding Therefor.

TITLE OF INVENTION: Methods of

PRODUCTION AND USE

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

ADDRESSEE: Felie & Lynch

STREET: 805 Third Avenue

CITY: New York City

STATE: New York

COUNTRY: USA

ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 360 kb storage

COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/576.775A

FILING DATE: 21-December-1995

CLASSIFICATION: 514

Prior APPLICATION DATA:

APPLICATION NUMBER: 08/ 3,133

FILING DATE: 17-July-1994

Prior APPLICATION DATA:

APPLICATION NUMBER: PC 94/04289

FILING DATE: 22-December-1994

ATTORNEY/AGENT INFORMATION:

NAME: Hanson, No. 584726man D.

REGISTRATION NUMBER: 0,946

REFERENCE/DOCKET NUMBER: BOER 1,500

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 688-9200

TELEFAX: (212) 688-9200

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 359 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE 359 AA: 41227 MW: 690704 CN:

Query Match 63.9% Score 46: DB 2: Length 359;

Best Local Similarity 75.0% Pred. No. 1:12e-02;

Matches 6: Conservative 2: Mismatches 0: Indels 0: Gaps 0:

DB 348 ALKLTGK 355

QY 2 SLKLTSGK 9

RESULT 13

ID US-08-899-545-2 STANDARD: PRT: 359 AA:

XX AC xxxxxx

DI

DE

XX

XX

W P O R E A
***** (TW)

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MFASTCH_PP protein - protein database search, using Smith-Waterman algorithm
Run on: Mon Jun 19 16:23:50 2000; MasPar time 14.77 seconds
Tabular output not generated. 81.852 Million cell updates/sec

Title: >US-09-142-524A-6
Description: (1-12) from US09142524A.pep
Perfect Score: 72
Sequence: 1 ISLKTSGKIAS 12

Scoring table: PAM 150
Gap 15

Searched: 72,1208 seqs, 100765575 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-repding
1: PCT 2:06 3:060 4:07 5:080 6:051 7:082 8:083 9:084A
10:084B 11:085 12:086 13:087 14:088 15:089 16:090 17:091
18:092 19:093 20:094 21:095 22:NEWP 23:NEW50 24:NEW08
25:NEW09

Statistics: Mean 18.571; Variance 42.966; scale 0.432

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

| SUMMARIES | | | | |
|-----------|-------|-------|--------|-----------------------------------|
| Result | Query | Match | Length | Pred. No. |
| No. | Score | DB | ID | Description |
| 1 | 72 | 100.0 | 17 11 | US-08-526- Sequence 11, Applicati |
| 2 | 72 | 100.0 | 24 10 | US-08-467- Sequence 192, Applicat |
| 3 | 72 | 100.0 | 24 10 | US-08-467- Sequence 192, Applicat |
| 4 | 72 | 100.0 | 24 10 | US-08-468- Sequence 192, Applicat |
| 5 | 72 | 100.0 | 24 7 | US-08-226- Sequence 192, Applicat |
| 6 | 72 | 100.0 | 24 8 | US-08-350- Sequence 192, Applicat |
| 7 | 72 | 100.0 | 24 10 | US-08-467- Sequence 192, Applicat |
| 8 | 72 | 100.0 | 80 17 | US-08-142- Sequence 190, Applicat |
| 9 | 72 | 100.0 | 127 10 | US-08-467- Sequence 190, Applicat |
| 10 | 72 | 100.0 | 127 8 | US-08-350- Sequence 190, Applicat |
| 11 | 72 | 100.0 | 127 10 | US-08-467- Sequence 190, Applicat |
| 12 | 72 | 100.0 | 127 7 | US-08-226- Sequence 190, Applicat |
| 13 | 72 | 100.0 | 127 10 | US-08-467- Sequence 190, Applicat |
| 14 | 72 | 100.0 | 127 10 | US-08-468- Sequence 190, Applicat |
| 15 | 72 | 100.0 | 460 11 | US-08-571- Sequence 1, Applicatio |
| 16 | 72 | 100.0 | 460 15 | US-08-980- Sequence 1, Applicatio |
| 17 | 72 | 100.0 | 514 15 | US-08-980- Sequence 2, Applicatio |
| 18 | 72 | 100.0 | 514 4 | US-07-975- Sequence 2, Applicatio |
| 19 | 72 | 100.0 | 514 11 | US-08-526- Sequence 14, Applicati |
| 20 | 72 | 100.0 | 514 8 | US-08-350- Sequence 134, Applicat |

| | | | | |
|----|----|-------|--------|-------------------------------------|
| 21 | 72 | 100.0 | 514 10 | US-08-467- Sequence 134, Applicat |
| 22 | 72 | 100.0 | 514 10 | US-08-468- Sequence 134, Applicat |
| 23 | 72 | 100.0 | 514 11 | US-08-571- Sequence 134, Applicat |
| 24 | 72 | 100.0 | 514 10 | US-08-467- Sequence 134, Applicat |
| 25 | 72 | 100.0 | 514 10 | US-08-467- Sequence 134, Applicat |
| 26 | 72 | 100.0 | 514 8 | US-08-334- Sequence 2, Applicatio |
| 27 | 72 | 100.0 | 514 7 | US-08-226- Sequence 134, Applicat |
| 28 | 72 | 100.0 | 514 1 | PCT-US93-1 Sequence 2, Applicatio |
| 29 | 65 | 90.3 | 11 11 | US-08-526- Sequence 5, Applicatio |
| 30 | 65 | 90.3 | 17 11 | US-08-526- Sequence 22, Applicatio |
| 31 | 54 | 75.0 | 35 3 | US-60-172- Sequence 5, Applicatio |
| 32 | 52 | 72.2 | 74 3 | US-60-169- Sequence 4624, Applicat |
| 33 | 52 | 72.2 | 74 3 | US-60-160- Sequence 3284, Applicat |
| 34 | 52 | 72.2 | 100 3 | US-60-162- Sequence 2714, Applicat |
| 35 | 52 | 72.2 | 100 3 | US-60-169- Sequence 5348, Applicat |
| 36 | 50 | 59.4 | 199 20 | US-09-417- Sequence 25733, Applic |
| 37 | 49 | 58.1 | 298 20 | US-09-417- Sequence 38423, Applic |
| 38 | 49 | 58.1 | 394 17 | US-09-107- Sequence 5184, Applicat |
| 39 | 49 | 58.1 | 573 20 | US-09-450- Sequence 5090, Applicat |
| 40 | 49 | 58.1 | 1001 3 | US-60-173- Sequence 6781, Applicat |
| 41 | 49 | 58.1 | 1001 3 | US-60-167- Sequence 8461, Applicat |
| 42 | 48 | 56.7 | 77 3 | US-60-162- Sequence 2847, Applicat |
| 43 | 48 | 56.7 | 253 25 | US-07-452- Sequence 8265, Applicat |
| 44 | 48 | 56.7 | 318 20 | US-07-417- Sequence 25464, Applicat |
| 45 | 48 | 56.7 | 344 3 | US-60-128- Sequence 3273, Applicat |

ALIGNMENTS

RESULT 1
ID US-08-526-179-11 STANDARD: PRG: 17 AA.
XX
AC xxxxxx
XX
XX
XX
XX
DE Sequence 11, Application US/3526,179
XX
XX Sequence 11, Application US/PE 26179
CC GENERAL INFORMATION:
CC APPLICANT: SAITO, Saburo
CC APPLICANT: HINO, Katsuhiko
CC APPLICANT: TANIGUCHI, Yoshitomi
CC APPLICANT: KURIMOTO, Masashi
CC TITLE OF INVENTION: METHOD AND USES OF THE SAME
CC NUMBER OF SEQUENCES: 24
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: BROWDY AND NEINARK
CC STREET: 419 Seventh Street, N.W., Suite 300
CC CITY: Washington
CC STATE: D.C.
CC COUNTRY: USA
CC ZIP: 20004
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC Compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patentin Release #1.0, Version #1.30
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/526,179
CC FILING DATE:
CC CLASS/FICATION: 530
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: JP 242,137/1994
CC FILING DATE: 10-SEP-1994
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: JP 200,204/1995
CC FILING DATE: 14-JUL-1995
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: JP 200,221/1995
CC FILING DATE: 14-JUL-1995
CC ATTORNEY/AGENT INFORMATION:
CC NAME: BROWDY, Roger L.

CC REGISTRATION NUMBER: 25,618
CC REFERENCE/DOCKET NUMBER: SAITO-19
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: 202-628-5197
CC TELEFAX: 202-737-3528
CC TELEX: 248633
CC INFORMATION FOR SEQ ID NO: 11:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 17 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC SEQUENCE: 17 AA: 1778 MW: 1362 CN:

Query Match 100.0% Score 72: DB 11: Length 17:
Best Local Similarity 100.0% Pred. No. 1.63e-01:
Matches 12: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

CC 2 ISLKLTSKSIAS 13
CC 1 ISLKLTSKSIAS 12

RESULT 2
ID US-08-467-697-112 STANDARD: PRT: 24 AA.
XX XXXXX

Sequence 192, Application US/08467697

GENERAL INFORMATION:

CC APPLICANT: Griffith, Irwin J.
CC APPLICANT: Pollock, Joanne
CC APPLICANT: Bond, Julian F.
CC APPLICANT: Garman, Richard D.
CC APPLICANT: Kuo, Mei-Chang
CC APPLICANT: Yeung, Siu-mei H.
CC APPLICANT: Brauer, Andrew
CC APPLICANT: Exley, Mark A.
CC APPLICANT: Powers, Steven P.
CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
CC TITLE OF INVENTION: Japanese Cedar Pollen
CC NUMBER OF SEQUENCES: 261
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
CC STREET: 610 Lincoln St
CC CITY: Waltham
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02154
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/467,697
CC FILING DATE: June 6, 1995
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/350,225
CC FILING DATE: December 6, 1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Jane E. Remillard
CC REGISTRATION NUMBER: 38,872
CC REFERENCE/DOCKET NUMBER: 025.6 USD4 (IM-028CPD4)
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 227-7400
CC TELEFAX: (617) 227-5941

CC INFORMATION FOR SEQ ID NO: 192:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 24 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC FRAGMENT TYPE: Internal
CC SEQUENCE: 24 AA: 2551 MW: 3329 CN:

Query Match 100.0% Score 72: DB 10: Length 24:
Best Local Similarity 100.0% Pred. No. 1.61e-01:
Matches 12: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

CC 13 ISLKLTSKSIAS 24
CC 1 ISLKLTSKSIAS 12

RESULT 3
ID US-08-467-006-192 STANDARD: PRT: 24 AA.
XX XXXXX

Sequence 192, Application US/08467606

GENERAL INFORMATION:

CC APPLICANT: Griffith, Irwin J.
CC APPLICANT: Pollock, Joanne
CC APPLICANT: Bond, Julian F.
CC APPLICANT: Garman, Richard D.
CC APPLICANT: Kuo, Mei-Chang
CC APPLICANT: Yeung, Siu-mei H.
CC APPLICANT: Brauer, Andrew
CC APPLICANT: Exley, Mark A.
CC APPLICANT: Powers, Steven P.
CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
CC TITLE OF INVENTION: Japanese Cedar Pollen
CC NUMBER OF SEQUENCES: 261
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
CC STREET: 610 Lincoln St
CC CITY: Waltham
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02154
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/467,006
CC FILING DATE: June 6, 1995
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/350,225
CC FILING DATE: December 6, 1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Jane E. Remillard
CC REGISTRATION NUMBER: 38,872
CC REFERENCE/DOCKET NUMBER: 025.6 USD6 (IM-028CPD6)
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 227-7400
CC TELEFAX: (617) 227-5941
CC INFORMATION FOR SEQ ID NO: 192:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 24 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide

```
CC FRAGMENT TYPE: Internal
SQ SEQUENCE 24 AA: 2551 MW: 3329 CN:

Query Match 100.0% Score 72: DB 10: Length 24:
Best Local Similarity 100.0% Pred. No. 1.63e-01:
Matches 12: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Db 13 ISLKLTSKGIAS 24
    IIIIIIIIIII
QY 1 ISLKLTSKGIAS 12

RESULT 4
ID US-06-468-940-192 STANDARD: PRT: 24 AA.
XX AC xxxxxx
XX DT
XX DE
XX SE Sequence 192, Application US/08468940
XX SE Sequence 192, Application US/08468940
CC GENERAL INFORMATION:
CC APPLICANT: Griffith, Irwin J.:
CC APPLICANT: Pollock, Joanne:
CC APPLICANT: Bond, Julian F.:
CC APPLICANT: Garman, Richard D.:
CC APPLICANT: Kuo, Mei-Chang:
CC APPLICANT: Yeung, Siu-mei H.:
CC APPLICANT: Brauer, Andrew:
CC APPLICANT: Exley, Mark A.:
CC APPLICANT: Powers, Steven P.:
CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
CC TITLE OF INVENTION: Japanese Cedar Pollen
CC NUMBER OF SEQUENCES: 261
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
CC STREET: 610 Lincoln St
CC CITY: Waltham
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02154
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Folio #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/468,940
CC FILING DATE:
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/350,225
CC FILING DATE: December 6, 1994
CC APPLICATION NUMBER: 08/226,248
CC FILING DATE: April 8, 1994
CC APPLICATION NUMBER: 07/938,990
CC FILING DATE: September 1, 1992
CC APPLICATION NUMBER: PCT/US93/00139
CC FILING DATE: January 15, 1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Darlene A. Vanstone
CC REGISTRATION NUMBER: 35,729
CC REFERENCE/DOCKET NUMBER: 025.6 US (IM1-028CP2)
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 466-6000
CC TELEFAX: (617) 466-6040
CC INFORMATION FOR SEQ ID NO: 192:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 24 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC FRAGMENT TYPE: Internal
SQ SEQUENCE 24 AA: 2551 MW: 3329 CN:

Query Match 100.0% Score 72: DB 7: Length 24:
Best Local Similarity 100.0% Pred. No. 1.63e-01:
Matches 12: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Db 13 ISLKLTSKGIAS 24
    IIIIIIIIIII
QY 1 ISLKLTSKGIAS 12

RESULT 5
ID US-08-226-248A-192 STANDARD: PRT: 24 AA.
XX AC xxxxxx
XX DT
XX DE
XX SE Sequence 192, Application US/08226248A
XX SE Sequence 192, Application US/08226248A
CC GENERAL INFORMATION:
CC APPLICANT: Griffith, Irwin J.:
CC APPLICANT: Pollock, Joanne:
CC APPLICANT: Bond, Julian F.:
CC APPLICANT: Garman, Richard D.:
CC APPLICANT: Kuo, Mei-Chang:
CC APPLICANT: Yeung, Siu-mei H.:
CC APPLICANT: Brauer, Andrew:
CC APPLICANT: Exley, Mark A.:
CC APPLICANT: Powers, Steven P.:
CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
CC TITLE OF INVENTION: Japanese Cedar Pollen
CC NUMBER OF SEQUENCES: 201
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
CC STREET: 610 Lincoln St
CC CITY: Waltham
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02154
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent In Folio #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/226,248A
CC FILING DATE: April 8, 1994
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/938,990
CC FILING DATE: September 1, 1992
CC APPLICATION NUMBER: PCT/US93/00139
CC FILING DATE: January 15, 1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Darlene A. Vanstone
CC REGISTRATION NUMBER: 35,729
CC REFERENCE/DOCKET NUMBER: 025.5 US (IM1-028CP)
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 466-6000
CC TELEFAX: (617) 466-6040
CC INFORMATION FOR SEQ ID NO: 192:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 24 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC FRAGMENT TYPE: Internal
SQ SEQUENCE 24 AA: 2551 MW: 3329 CN:

Query Match 100.0% Score 72: DB 7: Length 24:
Best Local Similarity 100.0% Pred. No. 1.63e-01:
Matches 12: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Db 13 ISLKLTSKGIAS 24
    IIIIIIIIIII
QY 1 ISLKLTSKGIAS 12
```



```

CC      ZIP: 02154
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC Compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DCS
CC      SOFTWARE: Patent-In Rel. use #1.0, Version: #1.25
CC      CURRENT APPLICATION DATA:
CC      APPLICATION NUMBER: US/39/467,066
CC      FILING DATE: June 6, 1995
CC      CLASSIFICATION: 424
CC      PRIOR APPLICATION DATA:
CC      APPLICATION NUMBER: 08/350,225
CC      FILING DATE: December 6, 1994
CC      ATTORNEY/AGENT INFORMATION:
CC      NAME: Jane E. remillard
CC      REGISTRATION NUMBER: 38,872
CC      REFERENCE/DOCKET NUMBER: 025,6 USD6 (IMI-028CPD6)
CC      TELECOMMUNICATION INFORMATION:
CC      TELEPHONE: (617) 227-7400
CC      TELEFAX: (617) 227-5941
CC      INFORMATION FOR SEQ ID NO: 190:
CC      SEQUENCE CHARACTERISTICS:
CC          LENGTH: 127 amino acids
CC          TYPE: amino acid
CC          TOPOLOGY: linear
CC      MOLECULE TYPE: peptide
CC      FRAGMENT TYPE: Internal
CC      SEQUENCE: 127 AA; 14435 MW; 84298 CN;
CC
CC      Query Match 100.0%; Score 72; aa 10; Length 127;
CC      Best Local Similarity 100.0%; Pred. No. 1; 6 Jc-Gj;
CC      Matches :2; Conservative 0; Mismatches 0; Indels 0;
CC
DB      11 ISLKTSGKIAS 22
CC      IIIIIIII
CC      I ISLKTSGKIAS 12
CC
RESULT 10
ID     US-08-350-225-190           STANDARD: PKI: 127 AA.
XX     xxxxxx
AC
AC
DE
DE
DE DE
XX     Sequence 190, Application US/39/467,066
CC      Sequence 190, Application US/08350225
CC      GENERAL INFORMATION:
CC      APPLICANT: Griffith, Irwin S.;
CC      APPLICANT: Pollock, Jeanne;
CC      APPLICANT: Bond, Julian P.;
CC      APPLICANT: Garman, Richard D.;
CC      APPLICANT: Kuo, Wei-Chang;
CC      APPLICANT: Yeung, Shu-mei H.;
CC      APPLICANT: Brauer, Andrew;
CC      APPLICANT: Exley, Mark A.;
CC      APPLICANT: Powers, Steven P.;
CC      TITLE OF INVENTION: Allergenic Proteins And Peptides From
CC      NUMBER OF SEQUENCES: 261
CC      CORRESPONDENCE ADDRESS:
CC      ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
CC      STREET: 610 Lincoln St.
CC      CITY: Waltham
CC      STATE: MA
CC      COUNTRY: USA
CC      ZIP: 02154
CC      COMPUTER READABLE FORM:
CC      MEDIUM TYPE: Floppy disk
CC      COMPUTER: IBM PC compatible
CC      OPERATING SYSTEM: PC-DOS/MS-DOS
CC      SOFTWARE: Patent-In Rel. use #1.0, Version: #1.25

```

CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/350.225
 CC FILING DATE: December 6, 1994
 CC CLASSIFICATION: 424
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/226.248
 CC FILING DATE: April 8, 1994
 CC ATTORNEY/AGENT INFORMATION:
 CC APPLICATION NUMBER: 07/938.990
 CC FILING DATE: September 1, 1992
 CC APPLICATION NUMBER: PC/US93/00139
 CC FILING DATE: January 15, 1993
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Carlene A. Vanstone
 CC REGISTRATION NUMBER: 35,729
 CC REFERENCE/DOCKET NUMBER: 025.6 US (.MI-028CP2)
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (517) 466-6090
 CC TELEFAX: (517) 466-6040
 CC INFORMATION FOR SEQ ID NO: 190:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 127 amino acids
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: peptide
 CC FRAGMENT TYPE: internal
 CC SEQUENCE 127 AA: 14435 MW: 84298 CN:
 CC
 CC Query Match 100.0% Score 72: DB 8: Length 127:
 CC Best Local Similarity 100.0% Pred. No. 1.63e-01:
 CC Matches 12: Conservative 0: Mismatches 0: Indels 0:
 CC
 CC DB 11 ISLKLTSKGIAS 22
 CC
 CC QY 1 ISLKLTSKGIAS 12
 CC
 CC RESULT 11
 CC ID US-08-467-023-190 STANDARD: PRT: 127 AA:
 CC XX xxxxxx
 CC
 CC DE Sequence 190: Application US/08/467023
 CC
 CC Sequence 190: Application US/08/467023
 CC GENERAL INFORMATION:
 CC APPLICANT: Griffith, Irwin J.
 CC APPLICANT: Pollock, Joanne
 CC APPLICANT: Bond, Julian F.
 CC APPLICANT: Garran, Richard D.
 CC APPLICANT: Kuo, Mei-Chang
 CC APPLICANT: Yeung, Siu-mei H.
 CC APPLICANT: Brauer, Andrew
 CC APPLICANT: Exley, Mark A.
 CC APPLICANT: Powers, Steven P.
 CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
 CC TITLE OF INVENTION: Japanese Cedar Pollen
 CC NUMBER OF SEQUENCES: 261
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
 CC STREET: 610 Lincoln St
 CC CITY: Waltham
 CC STATE: MA
 CC COUNTRY: USA
 CC ZIP: 02154
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patent In Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/467.023

CC FILING DATE: June 6, 1995
 CC CLASSIFICATION: 424
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 08/350.225
 CC FILING DATE: December 6, 1994
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Jane E. Remillard
 CC REGISTRATION NUMBER: 38,872
 CC REFERENCE/DOCKET NUMBER: 025.6 US2 (.MI-028CP2)
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (617) 227-7460
 CC TELEFAX: (617) 227-5941
 CC INFORMATION FOR SEQ ID NO: 190:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 127 amino acids
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: peptide
 CC FRAGMENT TYPE: internal
 CC SEQUENCE 127 AA: 14435 MW: 84298 CN:
 CC
 CC Query Match 100.0% Score 72: DB 10: Length 127:
 CC Best Local Similarity 100.0% Pred. No. 1.63e-01:
 CC Matches 12: Conservative 0: Mismatches 0: Indels 0:
 CC
 CC DB 11 ISLKLTSKGIAS 22
 CC
 CC QY 1 ISLKLTSKGIAS 12
 CC
 CC RESULT 12
 CC ID US-08-226-248A-190 STANDARD: PRT: 127 AA:
 CC XX xxxxxx
 CC
 CC DE Sequence 190: Application US/08/226248A
 CC
 CC Sequence 190: Application US/08/226248A
 CC GENERAL INFORMATION:
 CC APPLICANT: Griffith, Irwin J.
 CC APPLICANT: Pollock, Joanne
 CC APPLICANT: Bond, Julian F.
 CC APPLICANT: Garran, Richard D.
 CC APPLICANT: Kuo, Mei-Chang
 CC APPLICANT: Yeung, Siu-mei H.
 CC APPLICANT: Brauer, Andrew
 CC APPLICANT: Exley, Mark A.
 CC APPLICANT: Powers, Steven P.
 CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
 CC TITLE OF INVENTION: Japanese Cedar Pollen
 CC NUMBER OF SEQUENCES: 201
 CC CORRESPONDENCE ADDRESS:
 CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
 CC STREET: 610 Lincoln St
 CC CITY: Waltham
 CC STATE: MA
 CC COUNTRY: USA
 CC ZIP: 02154
 CC COMPUTER READABLE FORM:
 CC MEDIUM TYPE: Floppy disk
 CC COMPUTER: IBM PC compatible
 CC OPERATING SYSTEM: PC-DOS/MS-DOS
 CC SOFTWARE: Patent In Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/226.248A
 CC FILING DATE: April 8, 1994
 CC CLASSIFICATION: 424
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: 07/938.990
 CC FILING DATE: September 1, 1992
 CC APPLICATION NUMBER: PC/US93/00139

CC FILING DATE: January 15, 1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Darlene A. Vanstone
CC REGISTRATION NUMBER: 35,729
CC REFERENCE/DOCKET NUMBER: 025.5 US (IMI-028CP)
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 466-6000
CC TELEFAX: (617) 466-6040
CC INFORMATION FOR SEQ ID NO: 190:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 127 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC FRAGMENT TYPE: internal
CC SEQUENCE: 127 AA: 14435 MW: 84298 CN:

Query Match 100.0% Score 72: DB 7: Length 127:
Best Local Similarity 100.0% Pred. No. 1,53e-01:
Matches 12: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

DB 11 ISLKLTSKGIAS 22
QY 1 ISLKLTSKGIAS 12
|||||

RESULT 13
ID US-08-467-697-190 STANDARD: PRT: 127 AA:
XX xxxxxx
XX DT
XX DE
XX SEQUENCE 190. Application US/08457697
XX SEQUENCE 190. Application US/08467697
CC GENERAL INFORMATION:
CC APPLICANT: Griffith, Irwin J.
CC APPLICANT: Pollock, Joanne
CC APPLICANT: Bond, Julian F.
CC APPLICANT: Garman, Richard D.
CC APPLICANT: Kuo, Mei-Chang H.
CC APPLICANT: Yeung, Siu-mei H.
CC APPLICANT: Brauer, Andrew
CC APPLICANT: Exley, Mark A.
CC APPLICANT: Powers, Steven P.
CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
CC TITLE OF INVENTION: Japanese Cedar Pollen
CC NUMBER OF SEQUENCES: 251
CC CORRESPONDENCE ADDRESS:
CC ADDRESS: Immunologic Pharmaceutical Corporation, Inc.
CC STREET: 610 Lincoln St
CC CITY: Waltham
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02154
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/467,697
CC FILING DATE: June 6, 1995
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/350,225
CC FILING DATE: December 6, 1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Jane E. Reillard
CC REGISTRATION NUMBER: 38,872
CC REFERENCE/DOCKET NUMBER: 025.6 USD4 (IMI-028CPD4)
CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: (617) 227-7400
CC TELEFAX: (617) 227-5941
CC INFORMATION FOR SEQ ID NO: 190:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 127 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC FRAGMENT TYPE: internal
CC SEQUENCE: 127 AA: 14435 MW: 84298 CN:

Query Match 100.0% Score 72: DB 10: Length 127:
Best Local Similarity 100.0% Pred. No. 1,53e-01:
Matches 12: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

DB 11 ISLKLTSKGIAS 22
QY 1 ISLKLTSKGIAS 12
|||||

RESULT 14
ID US-08-468-940-190 STANDARD: PRT: 127 AA:
XX xxxxxx
XX DT
XX DE
XX SEQUENCE 190. Application US/08458940
XX SEQUENCE 190. Application US/08458940
CC GENERAL INFORMATION:
CC APPLICANT: Griffith, Irwin J.
CC APPLICANT: Pollock, Joanne
CC APPLICANT: Bond, Julian F.
CC APPLICANT: Garman, Richard D.
CC APPLICANT: Kuo, Mei-Chang H.
CC APPLICANT: Yeung, Siu-mei H.
CC APPLICANT: Brauer, Andrew
CC APPLICANT: Exley, Mark A.
CC APPLICANT: Powers, Steven P.
CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
CC TITLE OF INVENTION: Japanese Cedar Pollen
CC NUMBER OF SEQUENCES: 251
CC CORRESPONDENCE ADDRESS:
CC ADDRESS: Immunologic Pharmaceutical Corporation, Inc.
CC STREET: 610 Lincoln St
CC CITY: Waltham
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02154
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/458,940
CC FILING DATE:
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/350,225
CC FILING DATE: December 6, 1994
CC APPLICATION NUMBER: 08/226,248
CC FILING DATE: April 8, 1994
CC APPLICATION NUMBER: 07/938,990
CC FILING DATE: September 1, 1992
CC APPLICATION NUMBER: PCT/US93/00139
CC FILING DATE: January 15, 1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Darlene A. Vanstone
CC REGISTRATION NUMBER: 35,729
CC REFERENCE/DOCKET NUMBER: 025.6 US (IMI-028CP2)
CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: (617) 466-6000
CC TELEFAX: (617) 466-6040
CC INFORMATION FOR SEQ ID NO: 190:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 127 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC FRAGMENT TYPE: internal
CC SEQUENCE 127 AA: 14435 MW: 84258 CN:

Query Match 100.0% Score 72: DB 10: Length 127:
Best Local Similarity 100.0% Pred. No. 1.63e-01:
Matches 12: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Db 11 IS:KLTSGKIAS 22
QY 1 IS:KLTSGKIAS 12

RESULT 15
ID US-08-571-978-1 STANDARD: PRT: 460 AA
XX
AC xxxxxx
XX
CT
XX
DE

Sequence 1: Application US/C8571978

Sequence 1: Application US/08571978

GENERAL INFORMATION:

APPLICANT: KINO, KOISUKE

APPLICANT: KOMIYAMA, NAOKI

APPLICANT: SONE, TOSHIO

APPLICANT: KOHNO, YOICHI

TITLE OF INVENTION: ANTI-ALLERGIC AGENTS

NUMBER OF SEQUENCES: 36

CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,

STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400

CITY: ARLINGTON

STATE: VA

COUNTRY: USA

FIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk

COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/571,978

FILING DATE: 15-JAN-1996

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/JP94/01164

FILING DATE: 15-JUL-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 6-069336

FILING DATE: 07-APR-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 5-217725

FILING DATE: 01-SEP-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 5-177008

FILING DATE: 16-JUL-1993

ATTORNEY/AGENT INFORMATION:

NAME: OHLON, NORMAN F.

REGISTRATION NUMBER: 24,618

REFERENCE/SOCKET NUMBER: 7218-002-0 PCT

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703-413-3000

TELEFAX: 703-413-2220

CC INFORMATION FOR SEQ ID NO: 1:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 460 amino acids
CC TYPE: amino acid
CC STRANDEDNESS: single
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC SEQUENCE 460 AA: 50440 MW: 1058702 CN:

Query Match 100.0% Score 72: DB 11: Length 460:
Best Local Similarity 100.0% Pred. No. 1.63e-01:
Matches 12: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Db 344 IS:KLTSGKIAS 355
QY 1 IS:KLTSGKIAS 12

Search completed: Mon Jun 19 16:24:10 2000
Job time : 20 secs.


```
*****
WYEL (TM)
*****

Release A.L.A John F. Collins, Biocomputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Jun 19 16:04:40 2000: Maspar time 3.89 Seconds
145.422 Million cell updates/sec
Abular output not generated.
```

```
>US-09-i42-524A-6
(1-i2) from US09142524A.pcp
72
Perfect Score:
Sequence: 1 :SLKLTSGKIAS 12
```

ALIGNMENTS

```

RESULT 1
ENTRY
  TITLE second major allergen Cry j II precursor - Japanese cedar
  ORGANISM Formal Name Cryptomeria japonica common_name Japanese cedar
  DATE 16-Mar-1995 #sequence_revision 26-May-1995 #text_change
  12-Sep-1997

ACCESSIONS
  JC2498: PC2346; A60147

REFERENCE
  JC2498
  #authors Komiyama, N.; Sone, T.; Shimizu, K.; Morikubo, K.; Kino, K.
  #journal Biochem. Biophys. Res. Commun. (1994) 201:1021-1026
  #title cDNA cloning and expression of Cry j II, the second major
  allergen of Japanese cedar pollen.
  #cross-references MIM:194271186
  #accession JC2498
  #molecule_type mRNA
  #residues 1-514 #label COM
  #cross-references DDBJ:D297 ; NCBI:q556857 ; PIR:q506852
  #accession PC2346
  #molecule_type protein
  #residues 52-62 #label K02
  #reference A60147
  #authors Sakaguchi, M.; Inoue, S.; Takaki, M.; Ando, S.; Usui, M.;
  Matsumoto, T.
  #journal Allergy (1990) 45:309-312
  #title Identification of the second major allergen of Japanese cedar
  pollen.
  #cross-references MIM:190342988
  #accession A60147
  #molecule_type protein
  #residues 55-64 #label SAK
  #keywords glycoprotein; pollen

KEYWORDS
  FEATURE
  1-54
  55-460
  429,460,472
  #domain signal sequence #status predicted #label SIG
  #product second major allergen Cry j #status predicted
  #label MAT
  #binding site carbohydrate (Asn) (covalent) #status
  predicted
  SUMMARY
  #length 514 #molecular_weight 56583 #checksum 975
  Query Match 100.0%; Score 72; DB 2; Length 514;
  Best Local Similarity 100.0%; Pred. No. 1,816-05;
  Matches 12;..Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 398 ISLKTSGKIAS 409
Oy 1 ISLKTSGKIAS 12
IIIIIIIIII

```



```

Best Local Similarity 50.08: Pred. No. 2,9: 3-00;
Matches 6: Conservative 4: Mismatches 2: Indels 0: Gaps 0;

Db 21 VS:SLPTGKITA 32
    ||| |||
QY 1 IS:KLTSCKIA 12

RESULT 5
ENTRY T04273 #type complete
TITLE hypothetical protein F20B18.290 - Arabidopsis thaliana
ORGANISM #formal_name Arabidopsis thaliana #common_name mouse-ear
        cress
DATE 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change
        20-Sep-1999
ACCESSIONS T04273
REFERENCE T04273
#authors Bryan, M.; Rose, M.; Hempel, S.; Entian, K.D.; Hohetsel, J.;
#submission Newes, H.W.; Mayer, K.F.X.; Schueller, C.
#accession submitted to the Protein Sequence Database, March 1999
T04273
#molecule_type DNA
#residues 1-325 #label BEV
#cross-references EMBL:AF049483
#experimental_source cultivar Columbia; BAC clone F20B18
GENETICS
#map_position 4
#introns F20B18.290
#note #superfamily ADP,ATP carrier protein; ADP,ATP carrier protein
        repeat homology
CLASSIFICATION
#length 325 #molecular_weight 36551 #checksum 6140
SUMMARY
Query Match 58.18: Score 49: DB 2: Length 325:
Best Local Similarity 58.38: Pred. No. 7,948-00;
Matches 7: Conservative 2: Mismatches 3: Indels 0: Gaps 0;

Db 213 IS:KLVCOSVAG 224
    |||||
QY 1 IS:KLTSCKIA 12

RESULT 6
ENTRY E03409 #type complete
TITLE conserved hypothetical protein AF1278 - Archaeoglobus
        fulgidus
ORGANISM #formal_name Archaeoglobus fulgidus
DATE 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change
        05-Jun-1998
ACCESSIONS E03409
REFERENCE AK9250
#authors Kieck, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson,
        K.E.; Ketchum, K.A.; Dodson, R.J.; Quinn, M.; Hickey, E.K.;
        Peterson, J.D.; Richardson, D.L.; Kurlavage, A.R.; Graham,
        D.E.; Kyriakides, N.C.; Fleischmann, R.D.; Quackenbush, J.;
        Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;
        Dougherty, B.A.; McKenry, K.; Adams, M.D.; Loftis, B.;
        Peterson, S.; Reich, C.I.; McNeil, L.K.; Badger, J.H.;
        Glöck, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman,
        J.F.; McDonald, L.; Utterback, T.; Cotton, M.D.; Spriggs,
        T.; Artiach, P.; Kaine, B.P.; Sykes, S.M.; Sadow, P.W.;
        D'Andrea, K.P.; Bowman, C.; Fujita, C.; Garland, S.A.;
        Mason, T.M.; Olsen, G.J.; Fraser, C.M.; Smith, H.O.; Woese,
        C.R.; Venter, J.C.
#journal Nature (1997) 390:364-370
#title The complete genome sequence of the hyperthermophilic,
        sulfate-reducing archaeon Archaeoglobus fulgidus.
#cross-references M01D:98049343
#accession E03409
#status preliminary: nucleic acid sequence not shown;
        translation not shown
#molecule_type DNA
#residues 1-407 #label KLE

Best Local Similarity 50.08: Pred. No. 2,9: 3-00;
Matches 6: Conservative 4: Mismatches 2: Indels 0: Gaps 0;

```

```

**cross-references GB:AE001016: GB:AE000782: NID:q2589339: PID:q2649302:
TIGR:AF1278
SUMMARY #length 407 #molecular_weight 44931 #checksum 144
Query Match 68.18: Score 49: DB 2: Length 407:
Best Local Similarity 77.88: Pred. No. 7,948-00;
Matches 7: Conservative 1: Mismatches 1: Indels 0: Gaps 0;

Db 193 LKLGCKIA 201
    ||| |||
QY 3 LKLTSCIA 11

RESULT 7
ENTRY G70017 #type complete
TITLE N-carbamyl-L-amino acid amidohydrolase homolog yurH -
        Bacillus subtilis
ORGANISM #formal_name Bacillus subtilis
DATE 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change
        10-Sep-1999
ACCESSIONS G70017
REFERENCE A05580
#authors Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.;
        Alconi, G.; Azevedo, V.; Bertello, M.G.; Bessières, P.;
        Bototin, A.; Borchert, S.; Boriss, R.; Borriss, J.; Brans,
        A.; Braun, M.; Brigneau, S.C.; Bron, S.; Brovillat, S.;
        Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
        Choi, S.K.; Codanis, J.J.; Connerton, I.F.; Cummings, N.J.;
        Daniel, R.A.; Denizot, F.; Devine, K.M.; Duersthoefter, A.;
        Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Estlin, J.;
        Fabret, C.; Ferrari, E.; Foulger, D.; Fritz, C.; Fujita,
        M.; Fujita, Y.; Fuma, S.; Gallizi, A.; Galleron, N.; Gilm,
        S.Y.; Glaser, P.; Goffeau, A.; Golightly, E.J.; Grandi, G.;
        Guiseppi, G.; Guy, B.J.; Haga, K.; Halech, J.; Harwood,
        C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Husono, S.;
        Iulio, M.F.; Itaya, M.; Jones, L.; Joris, S.; Karamata, D.;
        Kasahara, Y.; Klier-Blanchard, M.; Klein, S.; Klamann, M.;
        Koetter, P.; Koningstein, G.; Krogh, S.; Kuma, M.;
        Kurita, K.; Lapidus, A.; Lardinois, S.; Lamber, J.;
        Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.;
        Mauvel, C.; Medigue, C.; Medina, N.; Mellado, R.P.; Mizuno,
        M.; Mostert, D.; Nakai, S.; Noback, M.; Noone, D.; O'Reilly,
        M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parry,
        V.; Pohl, E.M.; Portetle, P.; Porcelik, S.; Prestcott,
        A.M.; Prosser, E.; Pujic, P.; Purnelle, P.; Rappoport, S.;
        Rey, M.; Reynolds, S.; Rieger, M.; Rivolta, R.; Riba, E.;
        Kocet, B.; Rose, M.; Sadale, Y.; Sato, T.; Schmitt, E.;
        Schreier, S.; Schreier, M.; Scottone, F.; Sekiguchi, J.;
        Sekowska, A.; Serot, S.; Serron, P.; Shin, B.S.; Soldo,
        B.; Sorokin, A.; Tacconi, E.; Takaki, T.; Takahashi, H.;
        Takekura, K.; Takeuchi, M.; Tamakoshi, A.; Tanaka, T.;
        Terstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, S.;
        Vandenbol, M.; Vannier, F.; Vassarotti, A.; Viari, A.;
        Wambutt, R.; Wedler, E.; Wedler, H.; Weitzengraber, T.;
        Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto,
        K.; Yata, K.; Yoshida, K.; Yoshikawa, H.F.; Zimstein, E.;
        Yoshikawa, H.; Zanchin, A.
#journal Nature (1997) 390:249-256
#title The complete genome sequence of the Gram-positive bacterium
        Bacillus subtilis s.
#cross-references M01D:98044033
#accession G70017
#status preliminary: nucleic acid sequence not shown;
        translation not shown
#molecule_type DNA
#residues 1-412 #label KUN
#cross-references GB:P29120: GB:AL009126: NID:q2635613:
        PID:q2635750
        PID:q2635750
#experimental_source strain 168
GENETICS
#gene yurH
CLASSIFICATION
#superfamily N-carbamyl-L-amino acid amidohydrolase
#length 412 #molecular_weight 45519 #checksum 7655
SUMMARY

```

```

Query Match      68.1%   Score 49: DB 1: Length 412:
Best Local Similarity 60.0%: Pred. No. 7.94e-00:
Matches      6: Conservative      3: Mismatches      1: Indels      0: Gaps      0:

Db 263 LRLTCKGKIA 272
QY 3 LKLTSGKIAS 12

RESULT 9
ENTRY 105298
TITLE LKLTSGKIA 11
ALTERNATE_NAMES
ORGANISM Arabidopsis thaliana
DATE 23-Jul-1999
ACCESSIONS 105298
REFERENCE 215407
AUTHORS Boman, M.; Terry, N.; Ardiles, W.; Buyssebaert, C.;
Basseville, R.; De Clerck, R.; De Keyser, A.; Heyt, P.;
Rouze, P.; Van den Daele, H.; Villarroel, R.; Gielens, J.;
Van Montagu, M.; Hohn, J.; Mewes, H.W.; Mayer, K.F.X.;
Schueller, C.
SUBMISSION submitted to the Protein Sequence Database, October 1998
ACCESSION 105298
MOLECULE_TYPE DNA
RESIDUES 1-210
CROSS-REFERENCES EMBL:AL031804

Query Match      66.7%   Score 48: DB 2: Length 309:
Best Local Similarity 77.8%: Pred. No. 1.30e-01:
Matches      7: Conservative      1: Mismatches      1: Indels      0: Gaps      0:

Db 262 LKLTSGK 270
QY 1 LKLTSGK 9

RESULT 11
ENTRY S64479
TITLE probable membrane protein YGR169c - yeast (Saccharomyces cerevisiae)
ALTERNATE_NAMES
ORGANISM Saccharomyces cerevisiae
DATE 17-May-1996
ACCESSIONS S64479; S64482
REFERENCE S64071

**experimental_source cultivar Columbia: BAC clone F26p2:
GENETICS
  #map_position 4
  #clonings 14/3: 81/3: 102/1: 114/2
  #note F26p21.20
  #length 210 #molecular_weight 23517 #checksum 2724
SUMMARY
  Query Match      66.7%   Score 48: DB 2: Length 210:
  Best Local Similarity 70.0%: Pred. No. 1.30e-01:
  Matches      7: Conservative      2: Mismatches      1: Indels      0: Gaps      0:

Db 89 LKLTSGKIAS 98
QY 3 LKLTSGKIAS 12

RESULT 10
ENTRY A70194
TITLE ABC transporter, ATP-binding protein homolog - type disease Spirochete
ORGANISM Spirochete
DATE 13-Feb-1998
ACCESSIONS A70194
REFERENCE Fraser, C.M.; Caspers, S.; Blaud, W.M.; Sutton, G.; Clayton, R.; Lathigra, R.; White, O.; Ketchum, K.A.; Dodson, R.; Hickey, E.K.; Gwinn, M.; Dougherty, R.; Holt, J.F.; Fleischmann, R.D.; Richardson, D.; Peterson, J.; Karp, A.; Palmer, N.; Adams, M.D.; Gocayne, J.; White, O.; Utterback, T.; Wathey, L.; McDonald, L.; Artiach, P.; Bowman, C.; Garland, S.; Fujii, C.; Cotton, R.; Hest, K.; Roberts, K.; Hatch, R.; Smith, R.O.; Venter, J.C.; Nature (1997) 390:580-596
TITLE Genomic sequence of a type disease Spirochete, Borrelia burgdorferi
CROSS-REFERENCES MIM:19836594;
ACCESSION A70194
STATUS preliminary: nucleic acid sequence not shown: translation not shown
MOLECULE_TYPE DNA
RESIDUES 1-309
CROSS-REFERENCES GB:AB000750; NID:024496; PION:IA48038.1; PID:g21487
EXPERIMENTAL_SOURCE strain Desiree
CLASSIFICATION superfamily cytosolic aminopeptidase
KEYWORDS alpha-aminocyclopeptide hydrolase
SUMMARY
  Query Match      68.1%   Score 49: DB 1: Length 573:
  Best Local Similarity 45.5%: Pred. No. 7.94e-00:
  Matches      5: Conservative      5: Mismatches      1: Indels      0: Gaps      0:

Db 139 INRLPGGK 146
QY 1 LKLTSGKIA 11

RESULT 9
ENTRY 105298
TITLE LKLTSGKIA 11
ALTERNATE_NAMES
ORGANISM Arabidopsis thaliana
DATE 23-Jul-1999
ACCESSIONS 105298
REFERENCE 215407
AUTHORS Boman, M.; Terry, N.; Ardiles, W.; Buyssebaert, C.;
Basseville, R.; De Clerck, R.; De Keyser, A.; Heyt, P.;
Rouze, P.; Van den Daele, H.; Villarroel, R.; Gielens, J.;
Van Montagu, M.; Hohn, J.; Mewes, H.W.; Mayer, K.F.X.;
Schueller, C.
SUBMISSION submitted to the Protein Sequence Database, October 1998
ACCESSION 105298
MOLECULE_TYPE DNA
RESIDUES 1-210
CROSS-REFERENCES EMBL:AL031804

```

```

#authors Rieger, M.; Mueller-Auer, S.; Brueckner, M.; Schaefer, M.
#submission submitted to the Protein Sequence Database, May 1996
#accession S64479
##molecule_type DNA
##residues 1-376 ##label R1E
##cross-references EMBL:272953; NID:g1323295; PID:e243553; PID:g1323296;
MIPS:YGR168C
##experimental_source strain S288C
REFERENCE
S54003
#authors Hobling, U.; Hofmann, B.; Deilus, H.
#submission submitted to the Protein Sequence Database, May 1996
#accession S64482
##molecule_type DNA
##residues 1-376 ##label HEB
##cross-references EMBL:272953; NID:g1323295; PID:e243553; PID:g1323296;
MIPS:YGR168C
##experimental_source strain S288C
GENETICS
#map_position 78
#transmembrane protein
KEYWORDS
61-97
#domain transmembrane #status predicted #label TM1\
136-172
#domain transmembrane #status predicted #label TM2\
191-207
#domain transmembrane #status predicted #label TM3\
SUMMARY
#length 376 #molecular_weight 44301 #checksum 9227
Query Match 66.7%; Score 48; DB 2; Length 376;
Best Loca: Similarity 58.3%; Pred. No. 1.30e+01;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
DB 111 ISLMLTKGEVPS 122
1111111111
1 ISLKLTSKGKIAS 12
RESULT 12
ENTRY
#type complete
#hypothetical protein BB0261 - Lyme disease spirochete
#formal_name Borrelia burgdorferi #common_name Lyme disease
#spirochete
#10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change
10-Sep-1999
ACCESSIONS
E70132
REFERENCE
#authors
Fraser, C.M.; Castjens, S.; Huang, W.M.; Sutton, G.G.;
Clayton, R.; Lathigra, R.; White, O.; Ketchum, K.A.;
Duckson, R.; Hickey, E.K.; Gwinn, M.; Dougherty, B.; Tomb,
J.F.; Fleischmann, R.D.; Richardson, D.; Peterson, J.;
Kerlavage, A.R.; Quackenbush, J.; Salzberg, S.; Hanson, M.;
Vost, R.V.; Palmer, N.; Adams, M.D.; Gocayne, J.; Weidman,
J.; Uterback, T.; Watthey, L.; McDonald, L.; Artiach, P.;
Soman, C.; Garland, S.; Fujii, C.; Cotton, M.D.; Horst,
K.; Roberts, K.; Hatch, B.; Smith, R.C.; Venter, J.C.
#journal Nature (1997) 390:580-586
#title Genomic sequence of a Lyme disease spirochaete, Borrelia
burgdorferi.
#cross-references MJD:98065943
#accession E70132
#status preliminary; nucleic acid sequence not shown;
translation not shown
##molecule_type DNA
##residues 1-460 ##label KLE
##cross-references GB:AE00137; GB:AE000783; NID:g2688160; PID:g2688209;
TIGR:BB0261
#experimental_source strain B31
CLASSIFICATION
#superfamily hypothetical; protein BB0261; tetratricopeptide
repeat homology
FEATURE
103-136
#domain tetratricopeptide repeat homology #label TT1\
137-170
#domain tetratricopeptide repeat homology #label TT2\
205-238
#domain tetratricopeptide repeat homology #label TT3\
239-272
#domain tetratricopeptide repeat homology #label TT4\
278-310

```

```

311-344
SUMMARY
#atypical #label IT5\
#domain tetratricopeptide repeat homology #label IT6\
#length 460 #molecular_weight 54024 #checksum 9100
Query Match 66.7%; Score 48; DB 1; Length 460;
Best Loca: Similarity 77.8%; Pred. No. 1.30e+01;
Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
DB 23 IALKLTIGK 31
1111111111
1 ISKLTSKGK 9
QY
RESULT 13
ENTRY
#type complete
#major structural core protein - rice dwarf virus
#formal_name rice dwarf virus
#host Oryza sativa (rice)
#30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change
16-Jul-1999
ACCESSIONS
A45341; S12826; S12821
REFERENCE
A45341
#authors Suzuki, N.; Watanabe, Y.; Kusano, T.; Kitagawa, Y.
#journal Virology (1990) 179:455-459
#title Sequence analysis of the rice dwarf phyto-reovirus segment S3
transcript encoding for the major structural core protein
of 114 kDa.
#cross-references MJD:91021050
#accession A45341
##molecule_type mRNA
##residues 1-1019 ##label SUZ
##cross-references GB:X54520; NID:g61467; PID:CAA38440.1; PID:g61468
S12826
REFERENCE
S12826
#authors Yamada, N.; Uyeda, I.; Kudo, H.; Shikata, E.
#journal Nucleic Acids Res. (1990) 18:6419
#title Nucleotide sequence of rice dwarf virus genome segment 3.
#cross-references MJD:91057125
#accession S12826
##molecule_type genomic RNA
##residues 1-303; S:305-1019 ##label YAM
##cross-references EMBL:D00607; NID:g222500; PID:BAAC0482.1.
S12621
REFERENCE
S12621
#authors Kano, H.; Koizumi, M.; Noda, H.; Mizuno, H.; Tsukihara, T.;
Ishikawa, K.; Hibino, H.; Omura, T.
#journal Nucleic Acids Res. (1990) 18:6700
#title Nucleotide sequence of rice dwarf virus (RDV) genome segment
S3 coding for 114 k major core protein.
#cross-references MJD:91067474
#accession S12621
##molecule_type genomic RNA
##residues 1-18; Y:20-182; HG:185-245; Y:248-422; N:424-658.
#Y:660-809; N:811-909; N:911-1005; M:1007-1019
##label KAN
##cross-references GB:D00693
GENETICS
#map_position segment 3
CLASSIFICATION
#superfamily rice dwarf virus major structural core protein
#core protein
SUMMARY
#length 1019 #molecular_weight 114290 #checksum 2402
Query Match 66.7%; Score 48; DB 1; Length 1019;
Best Loca: Similarity 58.3%; Pred. No. 1.30e+01;
Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
DB 267 VSLKLTSGAGIGT 278
1111111111
1 ISKLTSKGKIAS 12
QY
RESULT 14
ENTRY
#type complete
#transcription-repair coupling factor (mfd) RP598 - Rickettsia

```

```

prowazekii
#organism #normal_name Rickettsia prowazekii
#date 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change
#accessions HT1564
#reference AD1530
#authors Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.;
Sicheritz-Ponten, T.; Alismark, U.C.M.; Podowski, R.M.;
Naeslund, A.K.; Eriksson, A.S.; Winkler, H.H.; Kurian,
C.G.
#journal Nature (1998) 396:133-140
#title The genome sequence of Rickettsia prowazekii and the origin
of mitochondria.
#cross-references MUID:99039499
#accession HT1564
#status preliminary: nucleic acid sequence not shown;
translation not shown
#molecule_type DNA
#residues 1-1120 #label AND
#cross-references GB:AJ235272; GB:AJ235269; NC:G3861033; PID:el342886;
PID:G3861142
#experimental_source strain Madrid E
GENETICS
#gene rpfJ: RP598
#classification #superfamily transcription-repair coupling protein: DEAD/H
box helicase homology
#length 1120 #molecular_weight 127674 #checksum 1852
SUMMARY
Query Match 66.7% Score 48; DB 2; Length 1120;
Best Local Similarity 70.0%; Pred. No. 1,30e-01;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
DB 675 IRLELASGK1 584
QY 1 ISLKLISGK1 10
RESULT 15
ENTRY
#title Prdage-related protein homolog yqbc - Bacillus subtilis
#organism Bacillus subtilis
#date 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change
#accessions B6948
#reference A69580
#authors Kist, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.;
Allard, G.; Azevedo, V.; Bortone, M.G.; Bessieres, P.;
Bollati, A.; Borcherdt, S.; Boriss, R.; Boursier, L.; Brans,
A.; Braun, M.; Briganti, S.C.; Bron, S.; Brouillet, S.;
Bruschi, C.V.; Caldwell, B.; Capriano, V.; Carter, N.M.;
Chen, S.K.; Codanis, J.J.; Connerton, I.F.; Cummings, N.J.;
Daniel, R.A.; Denizot, F.; Devine, K.M.; Duesterhoeft, A.;
Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.;
Fabret, C.; Ferrari, E.; Foulger, D.; Fritz, C.; Fujita,
M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galleron, N.; Ghim,
S.Y.; Glaser, P.; Goffeau, A.; Gollightly, E.J.; Grandi, G.;
Guiseppi, G.; Guy, B.J.; Haga, K.; Haiech, J.; Harwood,
C.R.; Henaut, A.; Hilbert, R.; Holsappel, S.; Hosono, S.;
Hullo, M.F.; Itaya, M.; Jones, L.; Joris, B.; Karamata, D.;
Kasahara, Y.; Klaerr-Blanchard, M.; Klein, C.; Kobayashi,
Y.; Koetter, P.; Koningstein, G.; Kroch, S.; Kumano, M.;
Kurita, K.; Lapidus, A.; Lardinois, S.; Lauber, J.;
Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.;
Maueel, C.; Medigue, C.; Medina, N.; Mellado, R.P.; Mizuno,
M.; Moestl, D.; Nakai, S.; Noback, M.; Noone, D.; O'Reilly,
M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro,
V.; Pohl, T.M.; Portetelle, D.; Porwolik, S.; Prescott,
A.M.; Prescan, E.; Pujic, P.; Purnelle, B.; Rapoport, G.;
Rey, M.; Reynolds, S.; Rieser, M.; Rivolta, C.; Rocha, E.;
Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon, E.;
Schleich, S.; Schroeter, R.; Sciffone, F.; Sekiguchi, J.;
Sekowska, A.; Seror, S.; Serror, P.; Shin, B.S.; Soldo,
B.; Sorckin, A.; Tacconi, E.; Takagi, T.; Takahashi, H.;

```

```

Takemaru, K.; Takeuchi, M.; Tamakoshi, A.; Tanaka, T.;
Terpstra, P.; Toqnoni, A.; Tosato, V.; Uchiyama, S.;
Vandenbol, M.; Vannier, F.; Vassarcotti, A.; Viari, A.;
Warbutt, R.; Wedler, E.; Wedler, H.; Weitzenegger, T.;
Winters, P.; Wips, A.; Yamamoto, H.; Yamane, K.; Yasunoto,
K.; Yata, K.; Yoshida, K.; Yoshikawa, H.F.; Zumbstein, E.;
Yoshikawa, H.; Zumbstein, A.
#journal Nature (1997) 390:249-256
#title The complete genome sequence of the gram-positive bacterium
Bacillus subtilis.
#cross-references MUID:98044333
#accession B6948
#status preliminary: nucleic acid sequence not shown;
translation not shown
#molecule_type DNA
#residues 1-1585 #label KUN
#cross-references GB:Z96117; GB:AA009126; NC:G2634966; PID:el193812;
PID:G2635048
#experimental_source strain 169
GENETICS
#gene yqbc
#summary #length 1585 #molecular_weight 171029 #checksum 3465
Query Match 66.7% Score 48; DB 2; Length 1585;
Best Local Similarity 70.0%; Pred. No. 1,30e-01;
Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
DB 326 LKLTGGQIAN 335
QY 3 LKLTSGKIAS 12

```

Search completed: Mon Jun 19 16:04:17 2000
Job time : 13 secs.

 W O R L D
 (TM)

Release 2.1A John F. Collins, BioComp, iting Research Unit,
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution rights by Oxford Molecular Ltd

Msrch_pp protein - protein database search, using Smith-Waterman algorithm
 Run on: Mon Jun 19 16:03:48 2000: MasPar time 2.83 Seconds
 129.156 Million cell updates/sec
 Total output not generated.

Title: US-09-142-524A-6
 Description: (1-12) from US09142524A.peg
 Perfect Score: 72
 Sequence: 1 ISKLISGXIAS 12

Scoring table: FAM 150
 Gap 15

Searched: 8,657 seqs. 30454573 residues

Post-processing: Minimum Match 0%
 Listing first 45 summaries

Database: swiss-prot38
 1:swissprot

Statistics: Mean 24.176; Variance 23.745; scale 1.018

Pred. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description | Pred. No. |
|------------|-------|-------------|--------|----|-----------------------------------|-----------|
| 1 | 72 | 100.0 | 514 | 1 | MPA2-CRYJA POSSIBLE POLYGALACTURO | 1.09e-06 |
| 2 | 53 | 73.6 | 316 | 1 | Y248-CAEEL HYPOTHETICAL 31.7 KD P | 2.36e-01 |
| 3 | 51 | 70.8 | 255 | 1 | FECE-ECOLI IRON(III) DICITRATE TR | 7.49e-01 |
| 4 | 49 | 68.1 | 573 | 1 | AMPL-SOLTC CHLOROPLAST AMINOPEPTI | 2.30e-00 |
| 5 | 48 | 66.7 | 376 | 1 | YGM-YEAST HYPOTHETICAL 44.3 KD P | 3.98e-00 |
| 6 | 48 | 66.7 | 1019 | 1 | VPI-RDV MAJOR 114 KD STRUCTURA | 3.98e-00 |
| 7 | 48 | 66.7 | 1120 | 1 | MFD-RICPR TRANSCRIPTION-REPAIR C | 3.98e-00 |
| 8 | 48 | 66.7 | 1585 | 1 | YQBC-BACSU HYPOTHETICAL 171.0 KD | 3.98e-00 |
| 9 | 47 | 65.3 | 530 | 1 | UL21-HSVB GENE 40 PROTEIN | 6.82e-00 |
| 10 | 47 | 65.3 | 572 | 1 | LAC3-THACU LACCASE 3 PRECURSOR (E | 6.82e-00 |
| 11 | 47 | 65.3 | 548 | 1 | PEX6-SCHPO PROBABLE PEROXIN-6 | 6.82e-00 |
| 12 | 46 | 63.9 | 121 | 1 | YF39-PYRHO HYPOTHETICAL PROTEIN P | 1.16e-01 |
| 13 | 46 | 63.9 | 124 | 1 | Y214-METJA PUTATIVE HYDROGENASE E | 1.16e-01 |
| 14 | 46 | 63.9 | 145 | 1 | AGGB-ECOLI AGGB PROTEIN PRECURSOR | 1.16e-01 |
| 15 | 46 | 63.9 | 208 | 1 | FTSQ-STGRG CELL DIVISION PROTEIN | 1.16e-01 |
| 16 | 46 | 63.9 | 251 | 1 | ZNUC-ECOLI HIGH-AFFINITY ZINC UPT | 1.16e-01 |
| 17 | 46 | 63.9 | 386 | 1 | BETA-ECOLI BETA-LACTAMASE PRECURS | 1.16e-01 |
| 18 | 46 | 63.9 | 316 | 1 | TALB-ECOLI TRANSALDOLASE B (EC 2 | 1.16e-01 |
| 19 | 46 | 63.9 | 359 | 1 | PST-CRIGR ALPHA-2,8-POLYSIALYLTR | 1.16e-01 |
| 20 | 46 | 63.9 | 359 | 1 | PST-HUMAN ALPHA-2,8-POLYSIALYLTR | 1.16e-01 |
| 21 | 46 | 63.9 | 359 | 1 | PST-MOUSE ALPHA-2,8-POLYSIALYLTR | 1.16e-01 |
| 22 | 46 | 63.9 | 551 | 1 | REFR-VARV RIFAMPICIN RESISTANCE | 1.16e-01 |
| 23 | 46 | 63.9 | 551 | 1 | REFR_VACCV RIFAMPICIN RESISTANCE | 1.16e-01 |

| | | | | | | |
|----|----|------|------|---|-----------------------------------|----------|
| 24 | 46 | 63.9 | 537 | 1 | NUSM-STRPV NADH-UBIQUINONE OXIDOR | 1.16e-01 |
| 25 | 46 | 63.9 | 836 | 1 | YU4K-CASEL HYPOTHETICAL 94.6 KD P | 1.16e-01 |
| 26 | 46 | 63.9 | 1010 | 1 | Y661-BASIN PROBABLE TONB-DEPENDEN | 1.16e-01 |
| 27 | 46 | 63.9 | 1202 | 1 | DPN-LASCIM PROBABLE DNA POLYMERAS | 1.16e-01 |
| 28 | 46 | 63.9 | 1361 | 1 | YML-YEAST HYPOTHETICAL 153.8 KD | 1.16e-01 |
| 29 | 45 | 62.5 | 97 | 1 | FLM-SALII NEGATIVE REGULATOR OF | 1.16e-01 |
| 30 | 45 | 62.5 | 97 | 1 | FLM-ECOLI NEGATIVE REGULATOR OF | 1.16e-01 |
| 31 | 45 | 62.5 | 330 | 1 | CAL-NAT CATHEPSIN S PRECURSOR | 1.16e-01 |
| 32 | 45 | 62.5 | 352 | 1 | Y456-METJA HYPOTHETICAL PROTEIN M | 1.16e-01 |
| 33 | 45 | 62.5 | 376 | 1 | KITH-HSV23 THYMIDINE KINASE (EC 2 | 1.16e-01 |
| 34 | 45 | 62.5 | 428 | 1 | RF4-KLUCA RF4 PROTEIN | 1.16e-01 |
| 35 | 45 | 62.5 | 463 | 1 | TIS-SALPO TYPE I RESTRICTION ENZ | 1.16e-01 |
| 36 | 45 | 62.5 | 556 | 1 | 2,3-BISPHOSPHOGLYCERAT | 1.16e-01 |
| 37 | 45 | 62.5 | 559 | 1 | PMGI-TOBAC 2,3-BISPHOSPHOGLYCERAT | 1.16e-01 |
| 38 | 45 | 62.5 | 638 | 1 | DNK-RHILE DNK PROTEIN (HEAT SHO | 1.16e-01 |
| 39 | 45 | 62.5 | 762 | 1 | SLAP-ACEMI CELL SURFACE PROTEIN P | 1.16e-01 |
| 40 | 45 | 62.5 | 1335 | 1 | XDH-DROME XANTHINE DEHYDROGENASE | 1.16e-01 |
| 41 | 44 | 61.1 | 269 | 1 | YRHF-ECOLI HYPOTHETICAL ABC TRANS | 3.23e-01 |
| 42 | 44 | 61.1 | 330 | 1 | CL-F-RAI CALPONIN, ACIDIC ISOFO | 3.23e-01 |
| 43 | 44 | 61.1 | 384 | 1 | AG-S-ECOLI AGAS PROTEIN | 3.23e-01 |
| 44 | 44 | 61.1 | 559 | 1 | YVI-BACSU HYPOTHETICAL 74.3 KD P | 3.23e-01 |
| 45 | 44 | 61.1 | 1850 | 1 | VI-2-CHICK VITELLOGENIN II PRECUR | 3.23e-01 |

ALIGNMENTS

| RESULT ID | MPA2-CRYJA | STANDARD: | RT: | 514 AA: |
|-----------|--|-----------|-----|---------|
| AC | P43212 | | | |
| DT | 01-NOV-1995 (Rel. 32, Created) | | | |
| DT | 01-NOV-1995 (Rel. 32, Last sequence update) | | | |
| DT | 01-NOV-1997 (Rel. 35, Last annotation update) | | | |
| DE | POSSIBLE POLYGALACTURONASE PRECURSOR (EC 3.2.1.15) (PG) (PECTINASE) | | | |
| DE | (MAJOR POLLEN ALLERGEN CRY J 2) (CRY J II) | | | |
| CS | Cryptomeria japonica (Japanese cedar) | | | |
| OC | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; | | | |
| OC | euphyllopytes; Spermatophyta; Coniferopsida; Coniferales; | | | |
| OC | Taxodiaceae; Cryptomeria | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. AND PARTIAL SEQUENCE | | | |
| RC | TISSUE-POLLEN | | | |
| RX | MEDLINE: 95010777 | | | |
| RA | Namba M., Kurose M., Torigoe K., Hino K., Taniguchi Y., Fukuda S. | | | |
| RA | Usui M., Kurimoto M. | | | |
| RT | Molecular cloning of the second major allergen, Cry J II, from | | | |
| RT | Japanese cedar pollen. | | | |
| RL | FEBS Lett. 353:124-128(1994). | | | |
| RN | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | TISSUE-POLLEN | | | |
| RX | MEDLINE: 94271186 | | | |
| RA | Komiyama N., Sone T., Shimizu K., Morikubo K., Kino K. | | | |
| RT | cDNA cloning and expression of Cry J II the second major allergen of | | | |
| RT | Japanese cedar pollen. | | | |
| RL | Biochem. Biophys. Res. Commun. 201:1021-1028(1994). | | | |
| RN | [3] | | | |
| RP | SEQUENCE OF 55-64 | | | |
| RX | MEDLINE: 90342988 | | | |
| RA | Sakaguchi M., Inouye S., Tanai M., Ando S., Usui M., Matukasi I. | | | |
| RT | Identification of the second major allergen of Japanese cedar | | | |
| RT | pollen. | | | |
| RL | Allergy 45:309-312(1990). | | | |
| CC | !- CATALYTIC ACTIVITY: RANDOM HYDROLYSIS OF 1,4-ALPHA-D- | | | |
| CC | GALACTOSIDURONIC LINKAGES IN PECTATE AND OTHER GALACTURONANS. | | | |
| CC | !- SUBCELLULAR LOCATION: SECRETED OR AMYLOPLAST (POTENTIAL). | | | |
| CC | !- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES | | | |
| CC | (POLYGALACTURONASES). | | | |

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial

TEM
PCP
BY

Commercial
Antibiotic/

01-OCT-1996 (Rel. 34, Last annotation update)
 DE CHLOROPLAST AMINOPEPTIDASE PRECURSOR (EC 3.4.11.1) (LEUCINE
 DE AMINOPEPTIDASE) (LAP) (LEUCYL AMINOPEPTIDASE) (PROLINE AMINOPEPTIDASE)
 DE (EC 3.4.11.5) (PROLYL AMINOPEPTIDASE).
 GN LAP.
 OS Solanum tuberosum (Potato).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC eupylophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 CC core eudicots; Asteridae; euasterids I; Solanales; Solanaceae;
 CC Solanum.
 RN [1]
 RF SEQUENCE FROM N.A.
 RC STRAIN=CV, DESIREE;
 RX MEDLINE: 94339796.
 RA Herbers K., Prat S., Willmitzer L.;
 RI Functional analysis of a leucine aminopeptidase from Solanum
 RI tuberosum L.;
 RI Planta 194:240-246(1994).
 RN [2]
 RP SEQUENCE OF 19-573 FROM N.A.
 RC STRAIN=CV, DESIREE; TISSUE=LEAF;
 RX MEDLINE: 93005746.
 RA Willmann T., Ebner M., Pena-Cortes H., Sanchez-Serrano J.J.,
 RA Willmitzer L., Prat S.;
 RI General roles of abscisic and jasmonic acids in gene activation: as a
 RI result of mechanical wounding.*;
 RI Plant Cell 4:1157-1170(1992).
 CC -!- FUNCTION: PRESUMABLY INVOLVED IN THE PROCESSING AND REGULAR
 CC TURNOVER OF INTRACELLULAR PROTEINS.
 CC -!- CATALYTIC ACTIVITY: RELEASE OF AN N-TERMINAL AMINO ACID, XAA-1-
 CC XRB-, IN WHICH XAA IS PREFERABLY LEU, BUT MAY BE OTHER AMINO ACIDS
 CC INCLUDING PRO ALTHOUGH NOT ARG OR LYS, AND XBB MAY BE PRO.
 CC -!- COFACTOR: BINDS TWO ZINC IONS (BY SIMILARITY).
 CC -!- SUBUNIT: HOMOHexamER (PROBABLE).
 CC -!- SUBCELLULAR LOCATION: CHLOROPLAST (BY SIMILARITY).
 CC -!- TISSUE SPECIFICITY: IN TUBERS AND FLORAL BUDS OF UNTREATED PLANTS.
 CC AFTER ABA TREATMENT OR MECHANICAL WOUNDING IS MOSTLY ACCUMULATED
 CC IN LEAVES, TO A LESSER EXTENT IN STEMS, BUT NOT IN ROOTS.
 CC -!- INDUCTION: BY ABSCISIC ACID (ABA); JASMONIC ACID (JA) AND
 CC WOUNDING.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M17; ALSO KNOWN AS THE
 CC CYTOSOL AMINOPEPTIDASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sit.ch/announce/>
 CC or send an email to license@isb-sit.ch).
 CC
 CC EMBL: X77219; CAA54334.1; .
 CC EMBL: X67845; CAA48038.1; .
 CC PIR: S24769; S24769.
 CC PIR: PQ0470; PQ0470.
 CC HSSP: PQ0727; ILAN.
 CC PRINTS: PR0481; LAMNOPPTIDASE.
 CC PROSITE: PS00631; CYTOSOLAP_1.
 CC PRAM: PF00883; Peptidase-M17_1.
 CC Transit peptide: Chloroplast; AminoPeptidase; Hydrolase; Zinc.
 CC TRANSIT 1 53 CHLOROPLAST AMINOPEPTIDASE.
 CC CHAIN 54 573 POLY-ALA.
 CC DOWN 169 174
 CC METAL 342 342 ZINC (2) (BY SIMILARITY).
 CC METAL 347 347 ZINC (1 AND 2) (BY SIMILARITY).
 CC METAL 357 367 ZINC (2) (BY SIMILARITY).
 CC METAL 427 427 ZINC (1) (BY SIMILARITY).
 CC METAL 429 429 ZINC (1 AND 2) (BY SIMILARITY).
 CC ACT_SITE 354 354 POTENTIAL.
 CC ACT_SITE 431 431 POTENTIAL.
 CC SEQUENCE 573 AA: 60122 MW: 31522454A47B291 CRG64;
 Query Match 68.1%; Score 49; DB 1; Length 573;
 RI kDa.";

Best Local Similarity 45.5%; Pred. No. 2,30e+00;
 Matches 5; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
 Db 139 INLRPGGRI 149
 QY : ISLKLISGKIA 11
 RESULT 5
 ID YG3W YEAST STANDARD: PRT: 376 AA.
 AC P53293;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE HYPOTHETICAL 44.3 KD PROTEIN IN CLC1-PDS2 INTERGENIC REGION.
 GN YGRI168C.
 OS Saccharomyces cerevisiae (Baker's Yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycetes; Saccharomycetales;
 CC Saccharomycetaceae; Saccharomycetes.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN= 288C;
 RX MEDLINE: 97435481.
 RA Rieger M., Bruecker M., Schaefer M., Mueller-Auer S.;
 RI Sequence analysis of 233 kilobases from Saccharomyces cerevisiae
 RI chromosome VII.*;
 RI Yeast 13:1077-1090(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Hebling J., Hofmann B., Delius H.;
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by on-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sit.ch/announce/>
 CC or send an email to license@isb-sit.ch).
 CC
 CC EMBL: Z72953; CAA97194.1; .
 CC KW Hypothetical protein; Transmembrane.
 CC TRANSMEM 153 173 POTENTIAL.
 CC TRANSMEM 188 208 POTENTIAL.
 CC SEQUENCE 376 AA: 44301 MW: 406940E805E8ADF CRG54;
 Query Match 60.7%; Score 49; DB 1; Length 376
 Best Local Similarity 56.4%; Pred. No. 3,96e+00;
 Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 Db 111 ISLKLISGKIPS 122
 QY : ISLKLISGKIAS 12
 RESULT 6
 ID VP3 RDV STANDARD: PRT: 1019 AA.
 AC P22472.
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-AUG-1991 (Rel. 19, Last sequence update)
 DT 01-MAR-1992 (Rel. 21, Last annotation update)
 DE MAJOR 114 KD STRUCTURAL PROTEIN.
 GN S3.
 OS Rice dwarf virus (RDV).
 CC Viruses; dsRNA viruses; Reoviridae; Fijivirus.
 RN [1]
 RP SEQUENCE FROM N.A.; AND SEQUENCE OF 501-519.
 RC STRAIN-ISOLATE A;
 RX MEDLINE: 91021050.
 RA Suzuki N., Watanabe Y., Kusano T., Kitagawa Y.;
 RI Sequence analysis of the rice dwarf phytoevirus segment S3
 RI transcript encoding for the major structural core protein of 114
 RI kDa.";

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL: C32216; BAA06447.1; .
 CC DR EMBL: D84432; BAA12411.1; .
 CC DF EMBL: 299117; CAB14544.1; .
 CC DR SUBMITTER: BG11285; YORO.
 CC DR PFAM: PF01464; SUT.1.
 CC KW Hypothetical protein.
 CC SQ SEQUENCE 1585 AA: 171030 MW: 1F82AB7FC250735F CRC64:

Query Match 66.7% Score 49; DB 1; Length 1585;

Best Local Similarity 70.0%; Pred. No. 3.98e+00;

Matches 7; Conservative 2; Mismatches 1; Indels 3; Gaps 0;

DL 325 LKLTSGQIAN 135

UY 3 LKLTSGKIAS 12

RESULT 9

ID UL21 HSVB STANDARD: PRT: 530 AA.

AC P28672

CI 01-DEC-1992 (Rel. 24, Created)

DI 01-DEC-1992 (Rel. 24, Last sequence update)

DE 01-DEC-1992 (Rel. 24, Last annotation update)

UN 40. GENE 40 PROTEIN.

CC Equine herpesvirus type 1 (strain Ab4p) (EHV-1).

CC Viruses: dsDNA viruses, no RNA stage; Herpesviridae;

CC Alphaherpesvirinae; Varicellovirus.

CC (1)

CC SEQUENCE FROM N.A.

CC MEDLINE: 92295566.

CC Telford E.A.R., Watson M.S., McBride K., Davison A.J.;

CC "The DNA sequence of equine herpesvirus-1";

CC Virology 189:304-316(1992)

CC -- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL21,
 CC EHV-1 4C, EHV-4 UL21, AND VZV 38.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL: M86654; AAB02476.1; .

CC DR PIR: F36799; W3BD4.

CC SQ SEQUENCE 530 AA: 57915 MW: 2E43565AB297CD18 CRC64:

Query Match

Best Local Similarity 65.3% Score 47; DB 1; Length 530;

Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

DB 126 VSKVTSSGLIVS 137

UY 1 LKLTSGKIAS 12

RESULT 10

ID LAC3_THACU STANDARD: PRT: 572 AA.

AC Q02079;

CI 01-NOV-1997 (Rel. 35, Created)

DI 01-NOV-1997 (Rel. 35, Last sequence update)

DE 01-NOV-1997 (Rel. 35, Last annotation update)

CC LACCASE 3 PRECURSOR (EC.1.10.3.2) (BENZENEDIOL-OXYGEN OXIDOREDUCTASE)

CC (URUSHIOL OXIDASE) (CATECHOL OXIDASE).

CC LCC3.

CC Thanatephorus cucumeris (Black scurf of potato) (Rhizoctonia solani).

CC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Ceratobasidiales;

Ceratobasidiaceae: anamorphic Ceratobasidiaceae; Rhizoctonia
 (1)
 RN SEQUENCE FROM N.A.
 RP STRAIN-RS22;
 RX MEDLINE: 96171523.
 RA Wahleithner J.A., Xu F., Brown K.M., Brown S.H., Gollightly E.J.,
 RA Hailer T., Kauppinen S., Pederson A., Schneider P.;

RT "The identification and characterization of four laccases from the
 RT plant pathogenic fungus Rhizoctonia solani";

CC Curr. Genet. 29:395-403(1996).

CC -- FUNCTION: LIGNIN DEGRADATION AND DETOXIFICATION OF LIGNIN-DERIVED
 CC PRODUCTS (PROBABLE).

CC -- CATALYTIC ACTIVITY: 4 BENZENEDIOL + O(2) -> 4 BENZOSEMIQUINONE +

CC 2 H2O.

CC -- COFACTOR: BINDS 4 CU-IONS PER MOLECULE, THREE DISTINCT CU

CC CENTERS KNOWN AS TYPE 1 OR BLUE, TYPE 2 OR NORMAL, AND TYPE

CC 3 OR COUPLED BINUCLEAR (BY SIMILARITY).

CC -- SUBUNIT: HOMODIMER (BY SIMILARITY).

CC -- SUBCELLULAR LOCATION: SECRETED.

CC -- TISSUE SPECIFICITY: IN MYCELIA, AT A LOWER LEVEL THAN LCC4.

CC -- SIMILARITY: BELONGS TO THE FAMILY OF MULTICOPPER OXIDASES.

CC -- SIMILARITY: CONTAINS 3 PLASTOCYANIN-LIKE DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL: Z54215; CAA90942.1; .

CC DR PROSITE: PS00079; MULTICOPPER_OXIDASE1; 1.

CC DR PROSITE: PS00080; MULTICOPPER_OXIDASE2; FALSE_NEG.

CC PFAM: PF00394; Cu-oxidase; 2.

CC Oxidoreductase; Signal; Copper Metal-binding; Glycoprotein; Repeat;

CC Lignin degradation; Multigene family; Polymorphism.

CC SIGNAL

CC CHAIN 19 572 "CASE 3."

CC FT DOMAIN 21 145 ASTOCYANIN-LIKE 1.

CC FT DOMAIN 157 304 ASTOCYANIN-LIKE 2.

CC FT DOMAIN 422 540 ASTOCYANIN-LIKE 3.

CC FT METAL 82 82 COPPER (TYPE 2) (BY SIMILARITY).

CC FT METAL 84 84 COPPER (TYPE 3) (BY SIMILARITY).

CC FT METAL 127 127 COPPER (TYPE 3) (BY SIMILARITY).

CC FT METAL 129 129 COPPER (TYPE 3) (BY SIMILARITY).

CC FT METAL 470 470 COPPER (TYPE 1) (BY SIMILARITY).

CC FT METAL 473 473 COPPER (TYPE 2) (BY SIMILARITY).

CC FT METAL 475 475 COPPER (TYPE 3) (BY SIMILARITY).

CC FT METAL 522 522 COPPER (TYPE 3) (BY SIMILARITY).

CC FT METAL 523 523 COPPER (TYPE 1) (BY SIMILARITY).

CC FT METAL 524 524 COPPER (TYPE 3) (BY SIMILARITY).

CC FT METAL 528 528 COPPER (TYPE 1) (BY SIMILARITY).

CC FT METAL 529 529 COPPER (TYPE 1) (BY SIMILARITY).

CC FT METAL 7 7 POTENTIAL.

CC FT CARBOHYD 182 182 POTENTIAL.

CC FT CARBOHYD 228 228 POTENTIAL.

CC FT CARBOHYD 294 294 POTENTIAL.

CC FT CARBOHYD 367 367 POTENTIAL.

CC FT CARBOHYD 405 405 POTENTIAL.

CC FT VARIANT 159 159 D -> N.

CC FT VARIANT 359 359 R -> H.

CC FT VARIANT 418 418 H -> Y.

CC FT VARIANT 448 448 I -> V.

CC SQ SEQUENCE 572 AA: 63747 MW: DF1E2AF9FF108CE4 CRC64:

Query Match

Best Local Similarity 65.3% Score 47; DB 1; Length 572;

Matches 6; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

DB 25 LKLTSGKIAIP 35

UY 2 LKLTSGKIAS 12

```

CC use by non-profit institutions as long as its content is in the way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC DR EMBL: AP000006; BAA30649.1;
CC KW Hypothetical protein.
CC SEQUENCE 121 AA: 13377 MW: 58124443.51988 CRC64:

Query Match 63.9% Score 46: DB 1: Length 124;
Best Local Similarity 60.0% Pred. No. 1:16e-01;
Matches 6: Conservative ? Mismatches 1: Indels 0: Gaps 0:

Db 17 LKLSGKIAA 26
| | | | |
QY 3 LKLSGKIAS 12
| | | | |

RESULT 13
ID Y214_METZ STANDARD: PPT: 124 AA.
AC Q57657
DI 01-NOV-1997 (Rel. 35: Created)
DI 01-NOV-1997 (Rel. 35: Last sequence update)
DI 01-NOV-1997 (Rel. 35: Last annotation update)
DE PUTATIVE HYDROGENASE EXPRESSION/FORMATION PROTEIN M00214
GN M00214
OS Methanococcus jannaschii
OC Archaea: Euryarchaeota: Methanococcales: Methanococcus
OC Methanococcus
OC [1]
RN SEQUENCE FROM N.A.
RP STRAIN: JAL-1 / DSM 2661 / ATCC 43067
RX MEDLINE: 96337999
RA But C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald C.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Iont J.-P., Adams M.D., Reishus C.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick C.M., Glouck A.,
RA Scott J.C., Geohagen N.M., Weidman J.F., Fainman J.L., Nguyen D.,
RA Uitterback T.R., Kelley J.M., Petersen J.D., Sedow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurt M.A., Kaine B.P., Horodovskiy M.,
RA Klek H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus
RA jannaschii."
RI Science 273:1058-1073(1996).
RL SIMILARITY: BELONGS TO THE HYPA/HYPE FAMILY.
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in the way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC DR EMBL: J67477; AAB98198.1;
CC TIGR: M0214
CC PROSITE: PS01249; HYPA: 1
CC PFAM: PF01155; HypA: 1
CC KW Hypothetical protein: Zinc-finger.
FT 2N-FING 78 100 C4-TYPE (POTENTIAL)
SQ SEQUENCE 124 AA: 14058 MW: 17698050.5496957AD CRC64:

Query Match 63.9% Score 46: DB 1: Length 124;
Best Local Similarity 66.7% Pred. No. 1:16e-01;
Matches 6: Conservative ? Mismatches 1: Indels 0: Gaps 0:

Db 103 IRKLSGGK 111
| | | | |
QY 1 ISKLSGK 9
| | | | |

RESULT 14
ID AGGB_ECOLI STANDARD: PPT: 145 AA.

```

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC -----
 DR EMBL: U07344; AAA56888.1; :
 KW Cell division.
 FT NON_TER
 FT DOMAIN <1 208
 SQ SEQUENCE 208 AA; 21942 MW; 7CC9AB6699746E82 CRC64;
 Query Match 63.9%; Score 45; DB 1; Length 208;
 Best Local Similarity 41.7%; Pred. No. 1; 156-01;
 Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
 DB 157 ISLRLTRDRVVT 168
 |||:| : :
 QY : ISLKLTSCKIAS 12
 Search completed: Mon Jun 19 16:03:54 2000
 Job time : 6 secs.

AC P46006;
 DI 01-NOV-1995 (Rel. 32, Created)
 DI 01-NOV-1995 (Rel. 32, Last sequence update)
 DI 01-NOV-1997 (Rel. 35, Last annotation update)
 DE AGGB PROTEIN PRECURSOR.
 SN AGGB.
 OS Escherichia coli.
 CC Plasmid P17-2.
 CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Escherichia.
 CC [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=ENTEROAGGREGATIVE 17-2 / SEROTYPE O3:H2;
 RX MEDLINE: 94327462.
 RA Savarino S.J., Fox P., Deng Y., Nataro J.P.;
 RT Identification and characterization of a gene cluster mediating
 RT enteroaggregative Escherichia coli aggregative adherence fimbria I
 R1 biogenesis.
 RL J. Bacteriol. 176:4949-4957(1994).
 CC - FUNCTION: NOT KNOWN.
 CC - SIMILARITY: HIGH, TO E. COLI AFAD PROTEIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC -----
 CC EMBL: U2894; AAA57453.1; :
 KW Fimbria; Plasmid; Signal.
 FT SIGNAL 1 24
 FT CHAIN 25 145 AGGB PROTEIN.
 SQ SEQUENCE 145 AA; 15837 MW; AD308CF0526F3DBF CRC64;
 Query Match 63.9%; Score 46; DB 1; Length 145;
 Best Local Similarity 50.0%; Pred. No. 1; 156-01;
 Matches 5; Conservative 5; Mismatch 0; Indels 0; Gaps 0;
 DB 42 GKMLATGRIA 51
 |||:| : :
 QY 2 SKLTSCKIA 11
 RESULT 15
 ID FTSQ_STRGR STANDARD; PRI: 208 AA.
 AC P45503;
 DI 01-NOV-1995 (Rel. 32, Created)
 DI 01-NOV-1995 (Rel. 32, Last sequence update)
 DE 01-NOV-1995 (Rel. 32, Last annotation update)
 DE CELL DIVISION PROTEIN FTSQ HOMOLOG (FRAGMENT).
 GN FTSQ.
 OS Streptomyces griseus.
 CC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 CC Actinomycetales; Streptomycetaceae; Streptomyces.
 CC [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=NRRL B2682;
 RX MEDLINE: 94374704.
 RA Dharmatilake A.J., Kendrick K.E.;
 RT Expression of the division-controlling gene *ftsZ* during growth and
 RT sporulation of the filamentous bacterium *Streptomyces griseus*.
 RL Gene 147:21-28(1994).
 CC - FUNCTION: THIS PROTEIN MAY BE INVOLVED IN ANOMALOUS FILAMENT
 CC GROWTH. MAY INTERACT WITH FTSSZ, FTSA, AND PBP3 (BY SIMILARITY).
 CC - SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN (BY SIMILARITY).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

| Result
No. | Score | Query | | DB | ID | Description | Prod. No. |
|---------------|-------|-------|--------|----|--------|------------------------|-----------|
| | | Match | Length | | | | |
| 1 | 53 | 73.6 | 316 | 5 | 076387 | C24G6.8 PROTEIN. | 5.98e-01 |
| 2 | 52 | 72.2 | 182 | 2 | 033358 | HYPOTHETICAL 19.4 KD P | 1.05e-00 |
| 3 | 51 | 70.8 | 623 | 10 | 092M11 | PUTATIVE RECEPTOR-LIKE | 1.83e-00 |
| 4 | 51 | 70.8 | 627 | 10 | 092M09 | PUTATIVE RECEPTOR-LIKE | 1.83e-00 |
| 5 | 50 | 69.4 | 457 | 2 | 092M85 | P2 ISBM1 | 3.17e-00 |
| 6 | 49 | 58.1 | 407 | 1 | 028990 | CONSERVED HYPOTHETICAL | 5.44e-00 |
| 7 | 49 | 58.1 | 412 | 2 | 032449 | YGRH PROTEIN. | 5.44e-00 |
| 8 | 48 | 66.7 | 210 | 0 | 082633 | HYPOTHETICAL 23.5 KD P | 9.27e-00 |
| 9 | 48 | 66.7 | 309 | 2 | 051595 | ABC TRANSPORTER, ATP-B | 9.27e-00 |
| 10 | 48 | 66.7 | 460 | 2 | 051276 | HYPOTHETICAL 54.0 KD P | 9.27e-00 |
| 11 | 48 | 66.7 | 1019 | 14 | 098630 | MAJOR CORE PROTEIN. | 9.27e-00 |
| 12 | 48 | 66.7 | 1743 | 5 | 019004 | SIMILAR TO GLUCAN 1. | 9.27e-00 |
| 13 | 47 | 65.3 | 199 | 1 | 028751 | CONSERVED HYPOTHETICAL | 1.57e+01 |
| 14 | 47 | 65.3 | 298 | 2 | 073352 | LACTOSE TRANSPORT SYST | 1.57e+01 |
| 15 | 47 | 65.3 | 332 | 5 | 093M83 | K04C1.6 PROTEIN. | 1.57e+01 |
| 16 | 47 | 65.3 | 381 | 14 | 031M85 | LDORF-1.2 PEPTIDE. | 1.57e+01 |
| 17 | 47 | 65.3 | 459 | 6 | 09X878 | T-CELL SURFACE GLYCOPR | 1.57e+01 |
| 18 | 47 | 65.3 | 525 | 4 | 093281 | COUNTERPART OF HSW-1 | 1.57e+01 |
| 19 | 47 | 65.3 | 594 | 2 | 092680 | POLYRIBONUCLEOTIDE NUC | 1.57e+01 |
| 20 | 47 | 65.3 | 907 | 11 | 092L14 | ORPHAN G PROTEIN-COUP | 1.57e+01 |

Query Match: 73.68; Score 53; DB 5; Length 316;
Best Local Similarity 58.38; P-red. NO. 5.98e-01;

Matches 7: Conservative 4: Mismatches 1: Indels 0: Gaps 0:

DE 209 MSKMGTKJAA 220
 QY 1 ISLKLTSKIAS 12

RESULT 2
 ID Q92W09 PRELIMINARY: PRT: 623 AA.
 AC Q92W09
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DI 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DE 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
 DE HYPOTHETICAL 19.4 KD PROTEIN.
 GN MTC260.0-033.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV.
 RA BROWN D., CHURCHER C.M.
 FL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV.
 RA PARKHILL J., BARRELL B.G., RAJANDREAM M.A.
 FL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-H37RV.
 RA MEDLINE; 96161548.
 RA PHILIPP W.J., POULET S., FICLMEIER K., PASCOPELLA L.,
 RA BALASUBRAMANIAN V., HEYM B., BERGH S., BLOOM B.R., JACOBS W.R. JR.,
 RA COLE S.L.
 FL "An integrated map of the genome of the tubercle bacillus.
 RI Mycobacterium tuberculosis H37RV, and comparison with Mycobacterium
 RI leprae."
 RC Proc. Natl. Acad. Sci. U.S.A. 94:3132-3137(1996).
 DR EMBL: Z97831; GAB0750.1;
 KW Hypothetical protein.
 SQ SEQUENCE 182 AA: 19401 MW: 25881D7F CRC32:

Query Match 72.2% Score 52: DB 2: Length 182:
 Best Local Similarity 63.6% Pred. No. 1.05e+00:
 Matches 7: Conservative 4: Mismatches 1: Indels 0: Gaps 0:

DE 131 ISLKLTSKIA 14:
 QY 1 ISLKLTSKIA 11

RESULT 3
 ID Q92W11 PRELIMINARY: PRT: 623 AA.
 AC Q92W11

DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DI 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DE 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
 DE PUTATIVE RECEPTOR-LIKE PROTEIN KINASE.
 GN F16P2.37.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC ephylllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
 OC Arabidopsis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA.
 RA ROUNSLEY S.D., LIN X., KAUL S., SHEA T.P., FUJII C.Y., MASON T.M.,
 RA SHEN M., RONNING C.M., FRASER C.M., SOMERVILLE C.R., VENTER J.C.;
 RI "Arabidopsis thaliana chromosome II BAC F16P2 genomic sequence."
 RI Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC004561; AAC95210.1;
 SQ SEQUENCE 182 AA: 19401 MW: 25881D7F CRC32:

DR HSSP: P02867; IRIN.
 KW Receptor; Kinase.
 SQ SEQUENCE 623 AA: 68216 MW: F0DAD5D7 CRC32:
 Query Match 70.8% Score 51: DB 10: Length 623:
 Best Local Similarity 77.8% Pred. No. 1.83e+00:
 Matches 7: Conservative 2: Mismatches 0: Indels 0: Gaps 0:

DE 185 ISLRLASOK 193
 QY 1 ISLKLISOK 9

RESULT 4
 ID Q92W09 PRELIMINARY: PRT: 627 AA.
 AC Q92W09
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DI 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
 DE PUTATIVE RECEPTOR-LIKE PROTEIN KINASE.
 GN F16P2.40.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC ephylllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 OC core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae;
 OC Arabidopsis.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA.
 RA ROUNSLEY S.D., LIN X., KAUL S., SHEA T.P., FUJII C.Y., MASON T.M.,
 RA SHEN M., RONNING C.M., FRASER C.M., SOMERVILLE C.R., VENTER J.C.;
 RI "Arabidopsis thaliana chromosome II BAC F16P2 genomic sequence."
 RI Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AC004561; AAC95213.1;
 DR HSSP: P05045; 15JQ.
 KW Receptor; Kinase.
 SQ SEQUENCE 627 AA: 68788 MW: 7060CA4D CRC32:

Query Match 70.8% Score 51: DB 10: Length 627:
 Best Local Similarity 77.8% Pred. No. 1.83e+00:
 Matches 7: Conservative 2: Mismatches 0: Indels 0: Gaps 0:

DE 185 ISLRLASOK 193
 QY 1 ISLKLISOK 9

RESULT 5
 ID Q92HX5 PRELIMINARY: PRT: 153 AA.
 AC Q92HX5
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DI 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
 DE P2 ISBM1.
 OS Brucella melitensis.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Brucella.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-16M;
 RA GODFROID F., TAMINIAU B., DANESI I., DENOEL P., TIBOR A., WEYNANTS V.,
 RA CLOECKAERT A., GODFROID J., LETESSON J.J.;
 RI "Identification of the perosamine synthetase gene of Brucella
 RI melitensis 16M and involvement of lipopolysaccharide O side chain in
 RI Brucella survival in mice and in macrophages."
 RI Infect. Immun. 66:0-0(1998).
 DR EMBL: AF047478; AAC58623.1;
 SQ SEQUENCE 157 AA: 17565 MW: 52393578 CRC32:

Query Match 69.4% Score 50: DB 2: Length 157:
 Best Local Similarity 54.5% Pred. No. 3.17e+00:
 Matches 5: Conservative 4: Mismatches 1: Indels 0: Gaps 0:

[illegible]

RA HILBERT H., HOLSAPPEL S., FORST S., HULLO M.F., ITAYA M., JONES L., JORIS B., KARAMATA C., KASAHARA Y., KLARR-BLANCHARD M., KLEIN C., KOBAYASHI Y., KOETTER P., KONSTEIN G., KROGH S., KUMANG M., KURITA K., LAPIDUS A., LARDIN S.S., LAUBER J., LAZAREVIC V., LEE S.M., LEVINE A., LIU H., MASUDA S., MAUEL C., MEDICUE C., MEDINA N., MELLADO R.P., MIZINGO M., MOESTL D., NAKAI S., NICKACK M., NOONE D., O'REILLY M., OGAWA K., OGIMAWA A., OUEDEGA B., PARK S.H., PARRO V., POHL T.M., PORTELELLA D., PORWOLLIK S., PRESCOTT A.M., PRESECAN E., PUJIC P., PURNELLE B., RAPAPORT G., REY M., REINOLDS S., RIEGER M., RIVOLTA C., ROCHA E., ROCHE B., ROSE M., SADAE Y., SATO T., SCANLAN E., SCHLECH S., SCHROETER R., SCOFFONE F., SEKIGUCHI J., SEKOWSKA A., SEPR S., SERROR P., SHIN B.S., SILDON H., SOKOKI A., TACCONE E., TAKAGI T., TAKAHASHI H., TAKEKAWA K., TAKEUCHI M., TAKAKOSHI A., TAKAKA T., TERPSTRA P., TONONI A., TOSATO V., UCHIYAMA S., VANDER L.M., VANNIER F., VASSAROTTI A., VIARI A., WAMBUIT R., WEDLER M., WEDLER H., WEITZNEGER T., WINTERS P., WIPAT A., YAMANO T., YAMANE K., YASUMOTO K., YATA K., YOSHIDA K., YOSHIKAWA H.F., YOSHIDA E., YOSHIMIZU H., YOSHIMIZU A., YOSHIMIZU K., YOSHIMIZU M., YOSHIMIZU N., YOSHIMIZU O., YOSHIMIZU P., YOSHIMIZU Q., YOSHIMIZU R., YOSHIMIZU S., YOSHIMIZU T., YOSHIMIZU U., YOSHIMIZU V., YOSHIMIZU W., YOSHIMIZU X., YOSHIMIZU Y., YOSHIMIZU Z., YOSHIMIZU AA, YOSHIMIZU AB, YOSHIMIZU AC, YOSHIMIZU AD, YOSHIMIZU AE, YOSHIMIZU AF, YOSHIMIZU AG, YOSHIMIZU AH, YOSHIMIZU AI, YOSHIMIZU AJ, YOSHIMIZU AK, YOSHIMIZU AL, YOSHIMIZU AM, YOSHIMIZU AN, YOSHIMIZU AO, YOSHIMIZU AP, YOSHIMIZU AQ, YOSHIMIZU AR, YOSHIMIZU AS, YOSHIMIZU AT, YOSHIMIZU AU, YOSHIMIZU AV, YOSHIMIZU AW, YOSHIMIZU AX, YOSHIMIZU AY, YOSHIMIZU AZ, YOSHIMIZU BA, YOSHIMIZU BB, YOSHIMIZU BC, YOSHIMIZU BD, YOSHIMIZU BE, YOSHIMIZU BF, YOSHIMIZU BG, YOSHIMIZU BH, YOSHIMIZU BI, YOSHIMIZU BJ, YOSHIMIZU BK, YOSHIMIZU BL, YOSHIMIZU BM, YOSHIMIZU BN, YOSHIMIZU BO, YOSHIMIZU BP, YOSHIMIZU BQ, YOSHIMIZU BR, YOSHIMIZU BS, YOSHIMIZU BT, YOSHIMIZU BU, YOSHIMIZU BV, YOSHIMIZU BW, YOSHIMIZU BX, YOSHIMIZU BY, YOSHIMIZU BZ, YOSHIMIZU CA, YOSHIMIZU CB, YOSHIMIZU CC, YOSHIMIZU CD, YOSHIMIZU CE, YOSHIMIZU CF, YOSHIMIZU CG, YOSHIMIZU CH, YOSHIMIZU CI, YOSHIMIZU CJ, YOSHIMIZU CK, YOSHIMIZU CL, YOSHIMIZU CM, YOSHIMIZU CN, YOSHIMIZU CO, YOSHIMIZU CP, YOSHIMIZU CQ, YOSHIMIZU CR, YOSHIMIZU CS, YOSHIMIZU CT, YOSHIMIZU CU, YOSHIMIZU CV, YOSHIMIZU CW, YOSHIMIZU CX, YOSHIMIZU CY, YOSHIMIZU CZ, YOSHIMIZU DA, YOSHIMIZU DB, YOSHIMIZU DC, YOSHIMIZU DD, YOSHIMIZU DE, YOSHIMIZU DF, YOSHIMIZU DG, YOSHIMIZU DH, YOSHIMIZU DI, YOSHIMIZU DJ, YOSHIMIZU DK, YOSHIMIZU DL, YOSHIMIZU DM, YOSHIMIZU DN, YOSHIMIZU DO, YOSHIMIZU DP, YOSHIMIZU DQ, YOSHIMIZU DR, YOSHIMIZU DS, YOSHIMIZU DT, YOSHIMIZU DU, YOSHIMIZU DV, YOSHIMIZU DW, YOSHIMIZU DX, YOSHIMIZU DY, YOSHIMIZU DZ, YOSHIMIZU EA, YOSHIMIZU EB, YOSHIMIZU EC, YOSHIMIZU ED, YOSHIMIZU EE, YOSHIMIZU EF, YOSHIMIZU EG, YOSHIMIZU EH, YOSHIMIZU EI, YOSHIMIZU EJ, YOSHIMIZU EK, YOSHIMIZU EL, YOSHIMIZU EM, YOSHIMIZU EN, YOSHIMIZU EO, YOSHIMIZU EP, YOSHIMIZU EQ, YOSHIMIZU ER, YOSHIMIZU ES, YOSHIMIZU ET, YOSHIMIZU EU, YOSHIMIZU EV, YOSHIMIZU EW, YOSHIMIZU EX, YOSHIMIZU EY, YOSHIMIZU EZ, YOSHIMIZU FA, YOSHIMIZU FB, YOSHIMIZU FC, YOSHIMIZU FD, YOSHIMIZU FE, YOSHIMIZU FF, YOSHIMIZU FG, YOSHIMIZU FH, YOSHIMIZU FI, YOSHIMIZU FJ, YOSHIMIZU FK, YOSHIMIZU FL, YOSHIMIZU FM, YOSHIMIZU FN, YOSHIMIZU FO, YOSHIMIZU FP, YOSHIMIZU FQ, YOSHIMIZU FR, YOSHIMIZU FS, YOSHIMIZU FT, YOSHIMIZU FU, YOSHIMIZU FV, YOSHIMIZU FW, YOSHIMIZU FX, YOSHIMIZU FY, YOSHIMIZU FZ, YOSHIMIZU GA, YOSHIMIZU GB, YOSHIMIZU GC, YOSHIMIZU GD, YOSHIMIZU GE, YOSHIMIZU GF, YOSHIMIZU GG, YOSHIMIZU GH, YOSHIMIZU GI, YOSHIMIZU GJ, YOSHIMIZU GK, YOSHIMIZU GL, YOSHIMIZU GM, YOSHIMIZU GN, YOSHIMIZU GO, YOSHIMIZU GP, YOSHIMIZU GQ, YOSHIMIZU GR, YOSHIMIZU GS, YOSHIMIZU GT, YOSHIMIZU GU, YOSHIMIZU GV, YOSHIMIZU GW, YOSHIMIZU GX, YOSHIMIZU GY, YOSHIMIZU GZ, YOSHIMIZU HA, YOSHIMIZU HB, YOSHIMIZU HC, YOSHIMIZU HD, YOSHIMIZU HE, YOSHIMIZU HF, YOSHIMIZU HG, YOSHIMIZU HH, YOSHIMIZU HI, YOSHIMIZU HJ, YOSHIMIZU HK, YOSHIMIZU HL, YOSHIMIZU HM, YOSHIMIZU HN, YOSHIMIZU HO, YOSHIMIZU HP, YOSHIMIZU HQ, YOSHIMIZU HR, YOSHIMIZU HS, YOSHIMIZU HT, YOSHIMIZU HU, YOSHIMIZU HV, YOSHIMIZU HW, YOSHIMIZU HX, YOSHIMIZU HY, YOSHIMIZU HZ, YOSHIMIZU IA, YOSHIMIZU IB, YOSHIMIZU IC, YOSHIMIZU ID, YOSHIMIZU IE, YOSHIMIZU IF, YOSHIMIZU IG, YOSHIMIZU IH, YOSHIMIZU II, YOSHIMIZU IJ, YOSHIMIZU IK, YOSHIMIZU IL, YOSHIMIZU IM, YOSHIMIZU IN, YOSHIMIZU IO, YOSHIMIZU IP, YOSHIMIZU IQ, YOSHIMIZU IR, YOSHIMIZU IS, YOSHIMIZU IT, YOSHIMIZU IU, YOSHIMIZU IV, YOSHIMIZU IW, YOSHIMIZU IX, YOSHIMIZU IY, YOSHIMIZU IZ, YOSHIMIZU JA, YOSHIMIZU JB, YOSHIMIZU JC, YOSHIMIZU JD, YOSHIMIZU JE, YOSHIMIZU JF, YOSHIMIZU JG, YOSHIMIZU JH, YOSHIMIZU JI, YOSHIMIZU JJ, YOSHIMIZU JK, YOSHIMIZU JL, YOSHIMIZU JM, YOSHIMIZU JN, YOSHIMIZU JO, YOSHIMIZU JP, YOSHIMIZU JQ, YOSHIMIZU JR, YOSHIMIZU JS, YOSHIMIZU JT, YOSHIMIZU JU, YOSHIMIZU JV, YOSHIMIZU JW, YOSHIMIZU JX, YOSHIMIZU JY, YOSHIMIZU JZ, YOSHIMIZU KA, YOSHIMIZU KB, YOSHIMIZU KC, YOSHIMIZU KD, YOSHIMIZU KE, YOSHIMIZU KF, YOSHIMIZU KG, YOSHIMIZU KH, YOSHIMIZU KI, YOSHIMIZU KJ, YOSHIMIZU KK, YOSHIMIZU KL, YOSHIMIZU KM, YOSHIMIZU KN, YOSHIMIZU KO, YOSHIMIZU KP, YOSHIMIZU KQ, YOSHIMIZU KR, YOSHIMIZU KS, YOSHIMIZU KT, YOSHIMIZU KU, YOSHIMIZU KV, YOSHIMIZU KW, YOSHIMIZU KY, YOSHIMIZU KZ, YOSHIMIZU LA, YOSHIMIZU LB, YOSHIMIZU LC, YOSHIMIZU LD, YOSHIMIZU LE, YOSHIMIZU LF, YOSHIMIZU LG, YOSHIMIZU LH, YOSHIMIZU LI, YOSHIMIZU LJ, YOSHIMIZU LK, YOSHIMIZU LL, YOSHIMIZU LM, YOSHIMIZU LN, YOSHIMIZU LO, YOSHIMIZU LP, YOSHIMIZU LQ, YOSHIMIZU LR, YOSHIMIZU LS, YOSHIMIZU LT, YOSHIMIZU LU, YOSHIMIZU LV, YOSHIMIZU LW, YOSHIMIZU LX, YOSHIMIZU LY, YOSHIMIZU LZ, YOSHIMIZU MA, YOSHIMIZU MB, YOSHIMIZU MC, YOSHIMIZU MD, YOSHIMIZU ME, YOSHIMIZU MF, YOSHIMIZU MG, YOSHIMIZU MH, YOSHIMIZU MI, YOSHIMIZU MJ, YOSHIMIZU MK, YOSHIMIZU ML, YOSHIMIZU MN, YOSHIMIZU MO, YOSHIMIZU MP, YOSHIMIZU MQ, YOSHIMIZU MR, YOSHIMIZU MS, YOSHIMIZU MT, YOSHIMIZU MU, YOSHIMIZU MV, YOSHIMIZU MW, YOSHIMIZU MX, YOSHIMIZU MY, YOSHIMIZU MZ, YOSHIMIZU NA, YOSHIMIZU NB, YOSHIMIZU NC, YOSHIMIZU ND, YOSHIMIZU NE, YOSHIMIZU NF, YOSHIMIZU NG, YOSHIMIZU NH, YOSHIMIZU NI, YOSHIMIZU NJ, YOSHIMIZU NK, YOSHIMIZU NL, YOSHIMIZU NM, YOSHIMIZU NO, YOSHIMIZU NP, YOSHIMIZU NQ, YOSHIMIZU NR, YOSHIMIZU NS, YOSHIMIZU NT, YOSHIMIZU NU, YOSHIMIZU NV, YOSHIMIZU NW, YOSHIMIZU NX, YOSHIMIZU NY, YOSHIMIZU NZ, YOSHIMIZU OA, YOSHIMIZU OB, YOSHIMIZU OC, YOSHIMIZU OD, YOSHIMIZU OE, YOSHIMIZU OF, YOSHIMIZU OG, YOSHIMIZU OH, YOSHIMIZU OI, YOSHIMIZU OJ, YOSHIMIZU OK, YOSHIMIZU OL, YOSHIMIZU OM, YOSHIMIZU ON, YOSHIMIZU OO, YOSHIMIZU OP, YOSHIMIZU OQ, YOSHIMIZU OR, YOSHIMIZU OS, YOSHIMIZU OT, YOSHIMIZU OU, YOSHIMIZU OV, YOSHIMIZU OW, YOSHIMIZU OX, YOSHIMIZU OY, YOSHIMIZU OZ, YOSHIMIZU PA, YOSHIMIZU PB, YOSHIMIZU PC, YOSHIMIZU PD, YOSHIMIZU PE, YOSHIMIZU PF, YOSHIMIZU PG, YOSHIMIZU PH, YOSHIMIZU PI, YOSHIMIZU PJ, YOSHIMIZU PK, YOSHIMIZU PL, YOSHIMIZU PM, YOSHIMIZU PN, YOSHIMIZU PO, YOSHIMIZU PP, YOSHIMIZU PQ, YOSHIMIZU PR, YOSHIMIZU PS, YOSHIMIZU PT, YOSHIMIZU PU, YOSHIMIZU PV, YOSHIMIZU PW, YOSHIMIZU PX, YOSHIMIZU PY, YOSHIMIZU PZ, YOSHIMIZU QA, YOSHIMIZU QB, YOSHIMIZU QC, YOSHIMIZU QD, YOSHIMIZU QE, YOSHIMIZU QF, YOSHIMIZU QG, YOSHIMIZU QH, YOSHIMIZU QI, YOSHIMIZU QJ, YOSHIMIZU QK, YOSHIMIZU QL, YOSHIMIZU QM, YOSHIMIZU QN, YOSHIMIZU QO, YOSHIMIZU QP, YOSHIMIZU QQ, YOSHIMIZU QR, YOSHIMIZU QS, YOSHIMIZU QT, YOSHIMIZU QU, YOSHIMIZU QV, YOSHIMIZU QW, YOSHIMIZU QX, YOSHIMIZU QY, YOSHIMIZU QZ, YOSHIMIZU RA, YOSHIMIZU RB, YOSHIMIZU RC, YOSHIMIZU RD, YOSHIMIZU RE, YOSHIMIZU RF, YOSHIMIZU RG, YOSHIMIZU RH, YOSHIMIZU RI, YOSHIMIZU RJ, YOSHIMIZU RK, YOSHIMIZU RL, YOSHIMIZU RM, YOSHIMIZU RN, YOSHIMIZU RO, YOSHIMIZU RP, YOSHIMIZU RQ, YOSHIMIZU RR, YOSHIMIZU RS, YOSHIMIZU RT, YOSHIMIZU RU, YOSHIMIZU RV, YOSHIMIZU RW, YOSHIMIZU RX, YOSHIMIZU RY, YOSHIMIZU RZ, YOSHIMIZU SA, YOSHIMIZU SB, YOSHIMIZU SC, YOSHIMIZU SD, YOSHIMIZU SE, YOSHIMIZU SF, YOSHIMIZU SG, YOSHIMIZU SH, YOSHIMIZU SI, YOSHIMIZU SJ, YOSHIMIZU SK, YOSHIMIZU SL, YOSHIMIZU SM, YOSHIMIZU SN, YOSHIMIZU SO, YOSHIMIZU SP, YOSHIMIZU SQ, YOSHIMIZU SR, YOSHIMIZ

Best Local Similarity 77.8%: Prod. No. 9.27e-00;
Matches 7: Conservative 1: Mismatches 1: Indels 0: Gaps 0:

DB 23 ISKLTSGK 31
QY 1 ISKLTSGK 9

RESULT 1:
ID Q98630 PRELIMINARY: PRT: 1019 AA.
AC Q98630:
DI 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DE 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE MAJOR CORE PROTEIN
DE RICE DWARF VIRUS (RDV)
DE Viruses: GRANA VIRUSES; Reoviridae; Flavivirus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CHINESE
RX MEDLINE: 91021050
RA SUZUKI N., WATANABE Y., KUSAN T., KITAGAWA Y.
RT "Sequence analysis of the rice dwarf phytochrome segment N-
terminal transcript encoding for the major structural core protein of the
RT Kda."
RT Viru.org 179:455-459(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CHINESE
RA ZHANG F., LIU Y., LIU Y., AN C., ZHENG Z.
RL Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.
DR EMBL: U72757; AAB17618.1;
SQ SEQUENCE 1019 AA: 114102 MW: 11100587 CRC12:

Query Match 66.7%: Score 48: DB 14: Length 1019;
Best Local Similarity 58.3%: Prod. No. 9.27e-00;
Matches 7: Conservative 4: Mismatches 1: Indels 0: Gaps 0:

DB 267 VSLKLTSGK 278
QY 1 ISKLTSGK 12

RESULT 12:
ID Q19004 PRELIMINARY: PRT: 1743 AA.
AC Q19004:
DI 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE 01-MAY-1999 (TrEMBLrel. 01, Last annotation update)
DE SMILAR TO GUCAN
CN D2096.3
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditidae.
OC Rhabditia; Rhabditidae; Pelodidae; Pelodidae.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 94150718
RA WILSON R., AINSWORTH R., ANDERSON K., BAYNES C., BERKS M.,
RA BONFIELD J., BURTON C., CONNELL M., COPSEY T., COOPER J., COULSON A.,
RA GRANTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
RA GARDNER A., GREEN P., HAKINS T., HILLIER L., JIER M., JOHNSTON L.,
RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
RA PARSONS J., PERCY C., RIEKEN L., ROOPRA A., SAUNDERS D., SHOWNKEEN R.,
RA SHALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
RA THIERRY-MIEG J., THOMAS K., VAUDIN M., VAUGHAN K., WATERSTON R.,
RA WATSON A., WEINSTOCK L., WILKINSON-SPROAT J., WOHLDMAN P.,
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans."
RL Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA GEISEL C.,

RESULT 9:
ID Q51695 PRELIMINARY: PRT: 309 AA.
AC Q51695:
DI 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DE 01-MAY-1999 (TrEMBLrel. 08, Last annotation update)
DE ABC TRANSPORTER, ATP-BINDING PROTEIN.
GN BB0754.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 35210 / H31;
RX MEDLINE: 98065943.
RA FRASER C.M., CASSENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,
RA LATHIGRA R., WHITE C., KETCHUM K.A., DODSON R., HICKEY E.K., GWINN M.,
RA DOUGHERTY B., TOMB J.-F., FLEISCHMANN R.D., RICHARDSON D.,
RA PETERSON J., KERLAVAGE A.R., QUACKENBUSH J., SALZBERG S.,
RA VAN VUOT R., PALMER N., ADAMS M.D., GOCAYNE J.D., WEIDMAN J.,
RA UTTERBACK T., WATTHEY L., McDONALD L., ARTIACH P., BOWMAN C.,
RA GARLAND S., FULI C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,
RA SMITH H.O., VENTER J.C.
RT "Genomic sequence of a Lyme disease spirochete, Borrelia
RT burgdorferi."
RL Nature 390:580-586(1997).
DR EMBL: AE001175; AAC67099.1;
RX TIGR: BB0754.
RP F1000005: ABC_tran: 1.
KW ATP-binding.
SQ SEQUENCE 309 AA: 45183 MW: 35566281 CRC12:

Query Match 96.7%: Score 48: DB 2: Length 309;
Best Local Similarity 77.8%: Prod. No. 9.27e-00;
Matches 7: Conservative 1: Mismatches 1: Indels 0: Gaps 0:

DB 262 ISKLTSGK 270
QY 1 ISKLTSGK 9

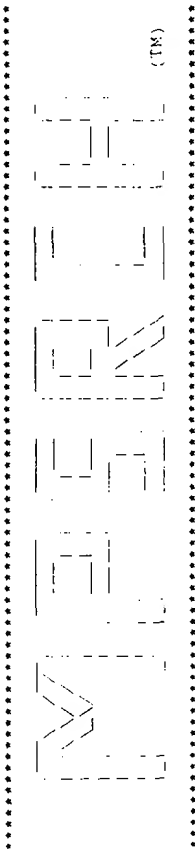
RESULT 10:
ID Q51276 PRELIMINARY: PRT: 460 AA.
AC Q51276:
DI 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DE 01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE HYPOTHETICAL 54.0 KD PROTEIN.
GN BB0261.
OS Borrelia burgdorferi (Lyme disease spirochete).
OC Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 35210 / H31;
RX MEDLINE: 98065943.
RA FRASER C.M., CASSENS S., HUANG W.M., SUTTON G.G., CLAYTON R.A.,
RA LATHIGRA R., WHITE C., KETCHUM K.A., DODSON R., HICKEY E.K., GWINN M.,
RA DOUGHERTY B., TOMB J.-F., FLEISCHMANN R.D., RICHARDSON D.,
RA PETERSON J., KERLAVAGE A.R., QUACKENBUSH J., SALZBERG S.,
RA VAN VUOT R., PALMER N., ADAMS M.D., GOCAYNE J.D., WEIDMAN J.,
RA UTTERBACK T., WATTHEY L., McDONALD L., ARTIACH P., BOWMAN C.,
RA GARLAND S., FULI C., COTTON M.D., HORST K., ROBERTS K., HATCH B.,
RA SMITH H.O., VENTER J.C.
RT "Genomic sequence of a Lyme disease spirochete, Borrelia
RT burgdorferi."
RL Nature 390:580-586(1997).
DR EMBL: AE001175; AAC66691.1;
RX TIGR: BB0261.
KW Hypothetical protein.
SQ SEQUENCE 460 AA: 54024 MW: 54143655 CRC12:

Query Match 56.7%: Score 48: DB 2: Length 460;

Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 ID 13
 AC SEQUENCE FROM N.A.
 RA WATERSTON R.
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U40800; AA08491.1
 DE PFAM: PF01055; Glyco_hydro_31.1
 DF PFAM: PF00888; trelol.1
 SQ SEQUENCE 1743 AA: 196517 MW: 196517 MW: C1EBB6BF CRC32:
 Query Match 55.7% Score 48: DB 5: Length 1743:
 Best Local Similarity 72.7% Pred. No. 9.27e-00:
 Matches 8: Conservative 2: Mismatches 1: Indels 0: Gaps 0:
 DB 944 SIKLTSQKIAS 954
 1111111111
 QY 2 SIKLTSQKIAS 12
 1111111111
 RESULT 13
 ID 028751 PRELIMINARY: PRI: 199 AA.
 AC 028751
 DT 01-JAN-1998 (FEBRELEI. 05, Created)
 DI 01-JAN-1998 (FEBRELEI. 05, Last sequence update)
 DE 01-NOV-1999 (FEBRELEI. 12, Last annotation update)
 DF CONSERVED HYDROPHILIC PROTEIN.
 DN AF1521
 OS Archaeoglobus fulgidus.
 NC Archaea: Euryarchaeota: Archaeoglobales: Archaeoglobaceae:
 OC Archaeoglobus.
 RN 111
 SEQUENCE FROM N.A.
 RA STRAIN-VC-15 / DSM 4304 / ATCC 49558.
 RX MEDLINE: 98045143
 PA KLECK H.P., CLAYTON R.A., TOMB J.F., WHITE O., NELSON K.E.,
 RA KEIGUM K.A., DOUSON R.J., GWINN M., HUCKEY E.K., PETERSON J.D.,
 RA PIRCHAPSON D.L., KEVLAVAGE A.R., GRAHAM D.E., KYRPIDES N.C.,
 RA FLEISCHMANN R.B., QUACKENBUSH J., LEE N.H., SUTTON G.G., GILL S.,
 RA KIRKNESS E.F., DOUGHERTY B.A., MCKENNEY K., ADAMS M.D., LOFTUS B.,
 RA PETERSON S., WEICH C.I., MCNEIL L.K., BADGER J.H., GLOBEK A., ZHOU L.,
 RA OVERBEEK R., SPRAGGS J.D., WEIDMAN J.F., McDONALD L., UTTERBACK T.,
 RA COTTON M.B., SPRAGGS T., ABIZACH P., KATNE B.P., SYKES S.M.,
 RA SADOW P.W., STANBREA K.P., BOWMAN C., FUJITA C., GARLAND S.A.,
 RA MASON T.M., OLSEN G.J., FRASER C.M., SMITH H.O., WOESE C.R.,
 RA VENTER J.C.
 OS The complete genome sequence of the hyperthermophilic sulphate-
 RE reducing archaeon Archaeoglobus fulgidus.
 RL Nature 390:364-370 (1997).
 DR EMBL: AEO00997; AAB89745.1
 DF TIGR: AF1521
 DP PFAM: PF01561; DUF27.1
 KW Hypothetical protein.
 SQ SEQUENCE 199 AA: 21856 MW: 59157A6: CRC32:
 Query Match 65.3% Score 47: DB 1: Length 199:
 Best Local Similarity 54.5% Pred. No. 1.57e-01:
 Matches 6: Conservative 3: Mismatches 2: Indels 0: Gaps 0:
 DB 19 SIKLQAQSDIT 29
 1111111111
 QY 1 SIKLTSQKIA 11
 1111111111
 RESULT 14
 ID P73352 PRELIMINARY: PRI: 298 AA.
 AC P73352
 DT 01-FEB-1997 (FEBRELEI. 02, Created)
 DI 01-FEB-1997 (FEBRELEI. 02, Last sequence update)
 DE 01-JAN-1999 (FEBRELEI. 09, Last annotation update)
 DF LACTOSE TRANSPORT SYSTEM PERMEASE PROTEIN LACF.
 DN LACF.
 OS Synchocystis sp. (strain PCC 6803).
 NC Bacteria: Cyanobacteria: Chroococcales: Synchocystis.
 OC

Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 ID 14
 AC SEQUENCE FROM N.A.
 RA STRAIN-PCC6803:
 RA TABATA S.
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U40800; AA08491.1
 DE PFAM: PF01055; Glyco_hydro_31.1
 DF PFAM: PF00888; trelol.1
 SQ SEQUENCE 1743 AA: 196517 MW: 196517 MW: C1EBB6BF CRC32:
 Query Match 55.7% Score 47: DB 2: Length 298:
 Best Local Similarity 41.7% Pred. No. 1.57e-01:
 Matches 5: Conservative 7: Mismatches 0: Indels 0: Gaps 0:
 DB 286 SIKLTSQKIAS 297
 1111111111
 QY 1 SIKLTSQKIAS 12
 1111111111
 RESULT 15
 ID 05KTH3 PRELIMINARY: PRI: 342 AA.
 AC 05KTH3
 DT 01-NOV-1999 (FEBRELEI. 12, Created)
 DI 01-NOV-1999 (FEBRELEI. 12, Last sequence update)
 DE 01-NOV-1999 (FEBRELEI. 12, Last annotation update)
 DF K04C1.6 PROTEIN
 DN K04C1.6
 OS Caenorhabditis elegans.
 NC Eukaryota: Metazoa: Nematoda: Secernentea: Rhabditida: Rhabditidae:
 OC Rhabditidae: Rhabditidae: Rhabditidae: Peloderidae: Caenorhabditis.
 RN 111
 SEQUENCE FROM N.A.
 RA MORTIMORE B.
 RX MEDLINE: 94150718.
 PA MORTIMORE B.
 OS *2.2 Mb of mitochondrial nucleotide sequence from chromosome III of C.
 RE elegans.
 RL Submitted (N.V.1996) to the EMBL/GenBank/DBJ databases.
 RA SEQUENCE FROM N.A.
 RX MEDLINE: 94150718.
 PA WILSON R., AINSKOUGH K., ANDERSON K., PAYNES C., BERKS M.,
 RA SUNFIELD J., BURTON J., CONNELL M., COPSEY T., COOPER J., CHOLIS N.A.,
 RA CRAXTON M., DEAR S., DU Z., DURBIN R., FAVELLO A., FULTON L.,
 RA GARDNER A., GREEN P., HARKINS T., HILLIER L., JIER M., JOHNSTON L.,
 RA JONES M., KERSHAW J., KIRSTEN J., LAISTER N., LATREILLE P.,
 RA LIGHTNING J., LLOYD C., MCMURRAY A., MORTIMORE B., O'CALLAGHAN M.,
 RA PARSONS J., PERCY C., RIEKEN L., ROOPRA A., SAUNDERS C., SHOWNKEN R.,
 RA SMALDON N., SMITH A., SONNHAMMER E., STADEN R., SULSTON J.,
 RA THIERFELDER M., THOMAS K., VAUGHAN K., VAUGHAN K., WATERSTON R.,
 RA WATSON A., WEINSTOCK L., WILKINSON-SPRAT J., WHILDEN P.
 RT *2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
 RE elegans.
 RL Nature 368:32-38 (1994).
 RN 131
 SEQUENCE FROM N.A.
 RC SPECIES=C elegans.
 RA AINSKOUGH R.
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 DR EMBL: Z81565; CAB04583.1
 DF EMBL: Z70267; CAB04583.1: JOINED.
 DR EMBL: Z70267; CAB04583.1

DR EMSL 2H1565: TAA44216.1: JOINED.
SQ SEQUENCE 332 AA: 34518 MW: 3DD15CED CRC32:
Query Match: 65.38: Score 47: DB 5: Length 332:
Best Local Similarity 45.58: Pred. No. 1.57e+01:
Matches 5: Conservative 5: Mismatches 1: Indels 0: Gaps 0:
DB 211:ILKUSGGRVP 221
QY 1:LSLKLISGKIA 11
Search completed: Mon Jun 19 16:04:24 2000
Job time : 13 secs.



Release 3.1A John F. Collins, BioComputing Research Unit,
Copyright (c) 1994-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

BLAST-pp protein - protein database search, using Smith-Waterman algorithm
Run on: Mon Jun 19 15:56:22 2000; MasPar time 5.66 Seconds
439,141 Million cell updates/sec
Partial output not generated.

Files: US-09-142-524A-2
Description: (1:05) from US09142524A.rag
Perfect Score: 406
Sequence: 1 MKVIVAFNFGPNRRVFVKR.....FDGRVDCGIIAAYONPASWK 105
Scoring table: BAW 150
Gap 11
Searched: 149961 seqs, 2356186 residues
Fast-processing: Minimum Match 08
Listing first 45 summaries
Database: a-geneseq35
lucaneseq
Statistics: Mean 29.704; Variance 127.75; scale 0.243

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

| SUMMARIES | | | |
|-----------|-----|-------|--------------|
| Result | No. | Score | Match length |
| 1 | 156 | 10.0 | 105 |
| 2 | 511 | 6.75 | 134 |
| 3 | 372 | 4.82 | 80 |
| 4 | 172 | 21.3 | 460 |
| 5 | 172 | 21.3 | 514 |
| 6 | 172 | 21.3 | 514 |
| 7 | 172 | 21.3 | 514 |
| 8 | 172 | 21.3 | 514 |
| 9 | 172 | 21.3 | 514 |
| 10 | 149 | 19.5 | 514 |
| 11 | 149 | 18.5 | 514 |
| 12 | 144 | 17.9 | 353 |
| 13 | 144 | 17.9 | 353 |
| 14 | 144 | 17.9 | 374 |
| 15 | 144 | 17.9 | 374 |
| 16 | 144 | 17.9 | 374 |
| 17 | 144 | 17.9 | 374 |
| 18 | 138 | 17.1 | 17 |
| 19 | 137 | 17.0 | 18 |
| 20 | 137 | 16.3 | 17 |
| 21 | 137 | 16.3 | 20 |
| 22 | 124 | 15.4 | 20 |
| 23 | 124 | 15.4 | 20 |

24 124 15.4 30 1 14579 Cry 1 1 pollen allergen 4 72e-02
25 123 15.3 47 1 14579 Sugi allergen protein 5 73e-02
26 117 14.5 14 1 1 535 Japanese cedar pollen 1 74e-01
27 117 14.5 15 1 1 497 Japan cedar pollen mat 1 74e-01
28 117 14.5 17 1 1 417 Sugi allergen protein 1 74e-01
29 115 14.3 17 1 1 4522 Cedar pollen allergen 2 51e-01
30 114 14.1 367 1 1 45573 Sugi s 1 2 62e-01
31 114 14.1 370 1 1 45578 Sugi s 1 2 62e-01
32 111 13.8 17 1 1 45580 Cedar pollen allergen 2 62e-01
33 111 13.8 17 1 1 45580 Sugi allergen protein 2 62e-01
34 110 13.6 13 1 1 45545 Japanese cedar pollen 2 62e-01
35 107 13.3 15 1 1 45768 Residues 66-80 of Cry 1 78e-00
36 107 13.3 15 1 1 45768 Japan cedar pollen mat 1 78e-00
37 107 13.3 47 1 1 45768 Sugi allergen protein 1 78e-00
38 106 13.2 13 1 1 45768 Japanese cedar pollen 1 78e-00
39 106 13.2 15 1 1 45768 Residues 76-90 of Cry 1 78e-00
40 106 13.2 15 1 1 45768 Japan cedar pollen mat 1 78e-00
41 106 13.2 20 1 1 45768 Cry 1 1 pollen allergen 1 78e-00
42 106 13.2 20 1 1 45768 Cry 1 1 Japanese cedar 1 78e-00
43 106 13.2 24 1 1 45768 Cry 1 1 pollen allergen 1 78e-00
44 106 13.2 30 1 1 45768 T-cell epitope peptide 1 78e-00
45 103 12.8 15 1 1 45768 Japan cedar pollen mat 2 21e-00

ALIGNMENTS

RESULT 1
ID W27370 standard: peptide: 105 AA
AC W27370
DI 24-MAR-1998 (first entry)
DE Multi-epitope peptide used as immunotherapeutic agent #2
KW Multi-epitope peptide; immunotherapeutic agent; allergic disease
XW T-cell epitope region; allergen; lymphocyte; immunodiagnosis
OS Synthetic
PN W09732600-A1
PO 12-SEP-1997
PF 10-MAR-1997 J00740
PR 10-MAR-1996 JP-080702
PA (MEIP) MELICI MILK PROD CO LTD
PI Dairiki K, Iwara A, Kino K, Kame A, Sone T
DR WPI: 97-470495/43
PT Peptide immunotherapeutic agent to treat allergic diseases
PT Contains multi-epitope peptide containing T cell epitope regions
PT from different allergens
PS Claim 6: Page 31: 58pp; Jan. 1990
CC The present sequence represents a multi-epitope peptide which is used as a new immunotherapeutic agent. It comprises T cell epitope regions from 2 or more different allergens (preferably linked via aromatic or lysine dimers), where the T cell epitope regions have a positive index greater than 100 as measured in a patient group responding to the allergen; have at least 75% reactivity with lymphocytes from patients responding to the allergen; and are not reactive with immunoglobulin E (IgE) antibodies from patients responsive to the allergen. The agent can be used to prevent and treat a wide variety of allergic diseases, e.g. by desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
SQ Sequence 105 AA:
Query Match: 100.0%; Score 806; DB 1; Length 105;
Best Local Similarity 100.0%; Pred. No. 2,14e-66;
Matches 105; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 MKVIVAFNFGPNRRVFVKRNV::HGRIDIFASKNFHLKNTIGTGRPKNNPILW 60
QY 1 MKVIVAFNFGPNRRVFVKRNV::HGRIDIFASKNFHLKNTIGTGRPKNNPILW 60
DB 61 FAKLTGFTLMGRRLKMPMYIAG::FDGRVDCGIIAAYONPASWK 105
QY 61 FAKLTGFTLMGRRLKMPMYIAG::FDGRVDCGIIAAYONPASWK 105

RESULT 2
ID W27371 standard: peptide: 134 AA
AC W27371

CC antibodies. 6 peptides (R81580-R81585) were identified as T-cell epitopes. These peptides, plus subsequences (R81573-75) essential for T-cell recognition, and homologous peptides (R81589-96) can be used as immunotherapeutic agents to treat or prevent cedar pollinosis, avoiding side-effects such as anaphylaxis.

Sequence 514 AA:

Query Match 21.8% Score 172 DB 1 Length 514
Best Local Similarity 95.2% Pred No. 4 23e-06
Matches 20 Conservative 1 Mismatches 0 Indels 0 Gaps 0

132 WKNNKINQCFALQTFENMG 152
||||| |||

51 WKNNKINQCFALQTFENMG 71

RESULT 13

ID W04346 standard: protein: 514 AA
AC W42122
DE 28-NOV-1998 (first entry)
KW Characyparis obtusa pollen allergen Cha o II
KW Pollen allergen Cha o II: T-cell epitope: prevention: treatment: pollenosis
KW Characyparis obtusa
OS Characyparis obtusa
PN J08176152-A1
PF 09-JUL-1996
PR 21-DEC-1994: J9589
PR 21-FEB-1994: JP-335089
PA (MEIP) MEIJ MILK PROD CO LTD.
PR WPI: 96-36825/37
PR N-PSUB: 1985+1
PI DNA encoding characyparis obtusa pollen allergen - T cell epitope(s) of which are useful in development of preventative and treating agent for Characyparis obtusa pollen pollinosis
PS Claim 14: Pages 14-15: 17pp: Japanese
TI The present sequence is the C. obtusa pollen allergen Cha o II, the T-cell epitopes of which can be used in the development of a preventative and treating agent for C. obtusa pollen pollinosis, C. obtusa pollen (2.4 kg) was decreased with dietary ether, and dried at room temp overnight. Cha o II was sepd. from it and purified. RNA was extracted from C. obtusa pollen, and cDNA and cDNA derived.

Sequence 514 AA:

Query Match 18.5% Score 143 DB 1 Length 514
Best Local Similarity 76.2% Pred No. 3 9e-04
Matches 17 Conservative 2 Mismatches 3 Indels 0 Gaps 0

132 WKNNKINQCFALQTFENMG 152

||||| |||

51 WKNNKINQCFALQTFENMG 71

RESULT 11

ID W42122 standard: protein: 514 AA
AC W42122
DE 16-JUN-1998 (first entry)
KW Japanese cypress pollen antigen Chao2
KW Japanese cypress pollen: antigen: T-cell epitope: Chao2: Chao2: diagnosis: allergy: spring tree pollen disease: pollinosis
OS Characyparis obtusa
PN W03747548-A1
PF 18-SEP-1997
PR 12-JUN-1997: J02031
PR 14-JUN-1996: JP-153527
PA (MEIP) MEIJ MILK PROD CO LTD.
PR Oat14: K. K. K. K.
PF WPI: 98-05242/35
TI T-cell epitope peptide portion of Japanese cypress pollen antigens Chao1 and Chao2 - used for diagnosis and treatment of spring tree pollen disease
Example 1. Page 15-18: 7pp: Japanese
TI The present sequence represents Japanese cypress pollen antigen Chao2.

CC The present invention describes peptides which correspond to the T-cell epitope sites on Japanese cypress pollen antigens Chao1 and Chao2. These peptides can be used as a reagent for the diagnosis of allergy to Japanese cypress pollen, and as an antigen in the treatment and prevention of spring tree pollen disease in which the pollinosis involves reactivity to Japanese cypress pollen.

Sequence 514 AA:

Query Match 18.5% Score 143 DB 1 Length 514
Best Local Similarity 76.2% Pred No. 3 9e-04
Matches 17 Conservative 2 Mismatches 3 Indels 0 Gaps 0

132 WKNNKINQCFALQTFENMG 152

||||| |||

51 WKNNKINQCFALQTFENMG 71

RESULT 12

ID R81587 standard: protein: 453 AA
AC R81587
DE 24-MAY-1995 (first entry)
KW Cedar pollen allergen B
KW Cedar pollen: allergen: immunoglobulin E: IgE: T-cell epitope: antibody: pollinosis: therapy: immunotherapy
OS Cryptomeria japonica
PN EP-700929-A2
PF 08-SEP-1995: J05295
PR 10-SEP-1994: JP-242137
PR 14-JUL-1995: JP-200221
PR 14-JUL-1995: JP-260204
PA (HAYB) HAYASHIBARA SEIBUTSU KAGAKU
PI Mito K, Saito S, Tanisuchi Y
DR WPI: 96-143976/15
TI New peptide(s) derived from cedar pollen allergens - active allergen-specific T-cells, but not allergen-specific IgE antibodies, used to treating cedar pollinosis
PS Claim 5: Page 31-32: 36pp: English
CC Synthetic peptides based on portions of cedar pollen antigens A (R81586) and B (R81587) were tested for their ability to activate cedar allergen-specific T-cells, but not allergen-specific IgE antibodies. 6 peptides (R81580-R81585) were identified as T-cell epitopes. These peptides, plus subsequences (R81573-75) essential for T-cell recognition, and homologous peptides (R81589-96) can be used as immunotherapeutic agents to treat or prevent cedar pollinosis, avoiding side-effects such as anaphylaxis.

Sequence 453 AA:

Query Match 17.9% Score 134 DB 1 Length 453
Best Local Similarity 85.0% Pred No. 1 0e-04
Matches 17 Conservative 1 Mismatches 2 Indels 0 Gaps 0

DB 75 MNIXLXPMYIAGYKTFDGR 95

||||| |||

70 MGRRLXPMYIAGYKTFDGR 89

RESULT 13

ID R75388 standard: protein: 353 AA
AC R75388
DE 12-MAR-1996 (first entry)
KW Japanese cedar pollen allergen Cry I
KW Japanese cedar: pollen allergen: Cry I: T-cell epitope: peptides: prevention: treatment: cryptoteria pollinosis
OS Cryptomeria japonica
PN Key
PF Location/Organism
TI peptide 51-75
TI /note="T-cell epitope peptide"
TI peptide 91-105
TI /note="T-cell epitope peptide"
TI peptide 106-123
TI /note="T-cell epitope peptide"
TI peptide 145-160

peptide /note= "T-cell epitope peptide"
 211..225
 peptide /note= "T-cell epitope peptide"
 326..340
 peptide /note= "T-cell epitope peptide"
 335..346
 peptide /note= "T-cell epitope peptide"
 307113295-A.
 PD 09-MAY-1995
 PR 20-OCT-1993: 2-2626.
 PR 20-OCT-1993: Jp-262626.
 PA (MEIJ) MEIJ MILK PROD CO LTD.
 LP WP1: 95-203834/21.
 PI New Cryptomeria pollen allergen T-cell epitope peptide - used for
 prevention, treatment and investigation of Japanese cedar pollinosis.
 PS Disclosures: Figs 1-2: 8pp: Japanese.
 CC R75384 is the Japanese cedar pollen allergen Cryj1, from which the
 T-cell epitope peptides R94289-R94295 were derived. The peptides
 can be used for the prevention and treatment of cryptomeria
 pollinosis, and also for the investigation of pollinosis.
 SQ Sequence: 354 AA.

Query Match 17.9% Score 144: DB 1: Length 353:
 Best Local Similarity 85.0% Pred. NO. 1.05e-03:
 Matches 17: Conservative 1: Mismatches 2: Indels 0: Gaps 0:

DE 75 MNKLKMPWYIAGYKTFDGR 95
 I I I I I I I I I I I
 CY 70 MGRRLKMPWYIAGYKTFDGR 89

RESULT 14
 ID: R50169 standard: Prote.in: 374 AA
 AC R50169
 DE 24-MAR-1995 (first entry)
 DE Japanese cedar pollen antigen Cryj1.
 KW Japanese cedar pollen antigen; Allergen; Cryj1; suai; pollinosis.
 OS Cryptomeria japonica.
 PE Key Location/Qualifiers
 FI peptide 1..21
 FI /label= signal_peptide
 FI 22..374
 FI /label= mature_Cryj1
 FN JC6197768-A.
 FI 19-OCT-1994
 FI 07-JAN-1993: 01115.
 PA (MEIJ) MEIJ SEIKA KAKISHA.
 LP WP1: 94-26268/23.
 LP N-PSDB: 071603.
 PI Sugi (Japanese cedar) pollen antigen Cryj1 - is used for
 diagnosis, treatment and prevention of sugi pollinosis.
 PS Claim 2: Page 5-7: 9pp: Japanese.
 CC The coding sequence for the Japanese cedar ("sugi") pollen allergen
 Cryj1 was isolated from a cDNA library prepared from polyA mRNA. All
 or part of the Cryj1 protein can be used for diagnosis, treatment
 and prevention of sugi pollinosis.
 SQ Sequence: 374 AA.

Query Match 17.9% Score 144: DB 1: Length 374:
 Best Local Similarity 85.0% Pred. NO. 1.05e-03:
 Matches 17: Conservative 1: Mismatches 2: Indels 0: Gaps 0:

DE 97 MNKLKMPWYIAGYKTFDGR 116
 I I I I I I I I I I I
 CY 70 MGRRLKMPWYIAGYKTFDGR 89

RESULT 15
 ID: R45541 standard: Prote.in: 374 AA.
 AC R45541
 DE 13-JUL-1994 (first entry)
 DE Cryj1: pollen allergen.

KW Japanese cedar: detection; allergy; treatment; diagnosis;
 OS T cell epitope; sensitivity.
 PE Cryptomeria japonica.
 FI Key Location/Qualifiers
 FI peptide 1..21
 FI /note= "signal peptide"
 FI 22..374
 FI /note= "mature peptide"
 FN W05401560-A.
 PD 20-JAN-1994
 PR 15-JAN-1993: U00139
 PR 10-JUL-1992: WC-U035661.
 PA (IMMUG) IMMUNOLOGIC PHARM CO
 PI Bond JF, Garman RD, Griffith LJ, Kuo M, Pollock J.
 DR WP1: 94-035056/04.
 DR N-PSDB: 055271.
 PI Antigens derived from Japanese cedar pollen allergen Cryj1
 PI contain at least two T cell epitopes, used to treat or diagnose
 PI allergy
 PS Disclosure: Fig 4: 137pp: English.
 CC The sequence is that of the Japanese cedar pollen allergen
 Cryj1 which contains at least two T cell epitopes. Peptide
 antigens derived from it can be used for the treatment and
 diagnosis of allergies associated with Japanese cedar pollen.
 CC The peptides have enhanced therapeutic properties but reduced
 CC side effects compared to naturally occurring allergens.
 SQ Sequence: 374 AA.

Query Match 17.9% Score 144: DB 1: Length 374:
 Best Local Similarity 85.0% Pred. NO. 1.05e-03:
 Matches 17: Conservative 1: Mismatches 2: Indels 0: Gaps 0:

DE 97 MNKLKMPWYIAGYKTFDGR 116
 I I I I I I I I I I I
 CY 70 MGRRLKMPWYIAGYKTFDGR 89

Search completed: Mon Jun 19 15:56:33 2000
 Job time: 11 secs.

[illegible]

20
 21
 22
 23
 24
 25
 26

(K1)

Release 3.1A John F. Collins, Biocomputing Research Unit,
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

protein - protein database search, using Smith-Waterman algorithm.

```
Run on: Mon Jun 19 15:59:53 2000; Master time 6.13 Seconds
405.568 Million cell updates/sec
```

Tabular output not generated

Title: UJS-09-142-524A-2
Description: (1-105) from US09142524A.pdf

Sequence: MKVTVAENQFGPNRRVFKE...FDGRRVGGTAAAYGNPASWK : 05

THE UNIVERSITY OF CHICAGO

searched : 06
sent : 90
received : 96

Post-processing: Minimum Match 08
 Listing first 45 samples
 Maximum Depth 105

base :
: geneseq.5
: geneseq.6

Statistik: Mean 29.704; Variance 127.355; scale 0.233

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARY

| Residue No. | Score | Query | | Length | DB ID | Description | Prod. No. |
|-------------|-------|-------|-----|--------|--------|-------------------------|-----------|
| | | Match | Gap | | | | |
| 1 | 406 | 100.0 | | 105 | W27370 | Multi-epitope peptide | 2.14e-65 |
| 2 | 372 | 46.2 | | 80 | W27369 | Multi-epitope peptide | 2.69e-24 |
| 3 | 138 | 17.1 | | 17 | R01582 | Cedar pollen allergen | 3.12e-03 |
| 4 | 137 | 17.0 | | 18 | W03346 | Sugi allergen protein | 4.02e-03 |
| 5 | 131 | 16.3 | | 17 | R01588 | Cedar pollen allergen | 1.26e-02 |
| 6 | 131 | 16.3 | | 20 | W42171 | T-cell epitope peptide | 1.26e-02 |
| 7 | 124 | 15.4 | | 20 | W45550 | Cry j I pollen allergen | 4.72e-02 |
| 8 | 124 | 15.4 | | 20 | R02459 | Cry j I Japanese cedar | 4.72e-02 |
| 9 | 124 | 15.4 | | 30 | R45579 | Cry j I pollen allergen | 4.72e-02 |
| 10 | 123 | 15.3 | | 47 | W80387 | Sugi allergen protein | 5.70e-02 |
| 11 | 117 | 14.5 | | 14 | W12535 | Japanese cedar pollen | 1.74e-01 |
| 12 | 117 | 14.5 | | 15 | R37887 | Japan cedar pollen mat | 1.74e-01 |
| 13 | 117 | 14.5 | | 17 | W03347 | Sugi allergen protein | 1.74e-01 |
| 14 | 115 | 14.3 | | 17 | R01582 | Cedar pollen allergen | 2.51e-01 |
| 15 | 111 | 13.8 | | 17 | R01580 | Cedar pollen allergen | 5.22e-01 |
| 16 | 111 | 13.8 | | 17 | W03345 | Sugi allergen protein | 5.22e-01 |
| 17 | 110 | 13.6 | | 13 | W12536 | Japanese cedar pollen | 6.57e-01 |
| 18 | 107 | 13.3 | | 15 | W57760 | Residues 66-80 of Cry | 1.08e-00 |
| 19 | 107 | 13.3 | | 15 | R37884 | Japan cedar pollen mat | 1.08e-00 |
| 20 | 107 | 13.3 | | 47 | W03356 | Sugi allergen protein | 1.08e-00 |
| 21 | 106 | 13.2 | | 13 | W12542 | Japanese cedar pollen | 1.29e-00 |
| 22 | 106 | 13.2 | | 15 | W57761 | Residues 76-90 of Cry | 1.29e-00 |

| | | | | | | | |
|----|-----|------|----|---|--------|-------------------------|---------|
| 23 | 106 | 13.2 | 15 | 1 | R9386 | Japan cedar pollen mat. | 1240-00 |
| 24 | 106 | 13.2 | 20 | 1 | R8549 | Cry 3 I pollen allerge | 1260-00 |
| 25 | 106 | 13.2 | 20 | 1 | R83478 | Cry 3 I Japanese Cedar | 1260-00 |
| 26 | 106 | 13.2 | 24 | 1 | R4580 | Cry 3 I pollen allerge | 1260-00 |
| 27 | 106 | 13.2 | 30 | 1 | R4582 | 1-cell epitope peptide | 1260-00 |
| 28 | 103 | 12.8 | 15 | 1 | F91859 | Japan cedar pollen mat | 1260-00 |

Note: post-processor removed 77 superfluous points from the data set.

APPENDICES

RESULTS 1

ID W27730 standard; peptide: 105 AA.

AC W27730;

DE 24-MAR-1998 (first entry)

DE Multi-epitope peptide used as immunotherapeutic agent #4.

KW Multi-epitope peptide: immunotherapeutic agent; allergic disease.

KW T-cell epitope region: allergen; lymphocyte; immunoglobulin E

OS Synthetic.

PN WC973260C-A1.

PO 12-SEP-1997.

PO 10-MAR-1997; J03740.

PP 10-MAR-1996; JP-080702.

PR (MEIP) MEIJI MILK PROD CO LTD.

PP Daiyiki K, Iwano K, Kume A, Some I.

DR WPI: 97-470495/43.

DR Peptide immunotherapeutic agent to treat allergic diseases .

PT Contains multi-epitope peptide containing T cell epitope regions

PT from different allergens

PS Claim 6; Page 31; 58pp; Japanese.

CC The present sequence represents a multi-epitope peptide which is used as

CC a new immunotherapeutic agent. It comprises T cell epitope regions from 2

CC or more different allergens (preferably linked via arginine or lysine

CC dimers), where the T cell epitope regions have a positivity index

CC greater than 100 as measured in a patient group responding to the

CC allergen; have at least 70% reactivity with lymphocytes from patients

CC responding to the allergen; and are not reactive with immunoglobulin E

CC (IgE) antibodies from patients responsive to the allergen; the agent can

CC be used to prevent and treat a wide variety of allergic diseases, e.g. by

CC desensitization. Side effects, e.g. those mediated by T cell, are reduced.

CC Sequence 105 AA:

CC

from different allergens
 Claim 5; Page 31; 58pp; Japanese.
 The present sequence represents a multi-epitope peptide which is used as a new immunotherapeutic agent. It comprises T cell epitope regions from 2 or more different allergens (preferably linked via arginine or lysine dimers), where the T cell epitope regions have a positivity index greater than 100 as measured in a patient group responding to the allergen; have at least 70% reactivity with lymphocytes from patients responding to the allergen; and are not reactive with immunoglobulin E (IgE) antibodies from patients responsive to the allergen. The agent can be used to prevent and treat a wide variety of allergic diseases, e.g. by desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
 Sequence 88 AA:

Query Match 46.2% Score 372; DB 1; Length 80;
 Best Local Similarity 100.0%; Pred. No. 2,096-24;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 MKVTVAFNFGNNRRVFKNVSNVHHRPRIDFASKNFHLOKNTIGTCRR 51
 IIII IIII IIII IIII IIII IIII IIII IIII IIII
 QY 1 MKVTVAFNFGNNRRVFKNVSNVHHRPRIDFASKNFHLOKNTIGTCRR 51

RESULT 3
 ID R81581 standard; Peptide: 17 AA.
 AC R81581

DE 24-MAY-1996 (first entry)
 DE Cedar pollen allergen peptide 9 (T-cell epitope).
 KW Cedar; Cryptomeria japonica; pollen; allergen; immunoglobulin E;
 KW IgE; T-cell; epitope; antibody; pollinosis; therapy; immunotherapy.
 OS Synthetic.
 PN EP-700429-A2
 PD 13-MAR-1996
 PF 08-SEP-1995; 406295.
 PR 10-SEP-1994; JP-242147.
 RR 14-JUL-1995; JP-200221.
 RR 14-JUL-1995; JP-200204.
 RA (HAYB.) HAYASHIBARA SEIBUTSU KAGAKU.
 RI Hiko K., Saito S., Taniguchi Y.
 DR WP: 96-140976/15.
 PT New peptide(s) derived from cedar pollen allergens - activate
 PT allergen-specific T-cells, but not allergen-specific IgE antibodies.
 PI used for treating cedar pollinosis
 IS Claim 4; Page 28; 36pp; English.
 CC Synthetic peptides based on portions of cedar pollen allergens A
 CC (R81586) and B (R81587) were tested for their ability to activate
 CC cedar allergen-specific T-cells, but not allergen-specific IgE
 CC antibodies. 6 peptides (R81582-R81587) were identified as T-cell
 CC epitopes. These peptides, plus subsequences (R81588-95) essential
 CC for T-cell recognition, and homologous peptides (R81589-95) can
 CC be used as immunotherapeutic agents to treat or prevent cedar
 CC pollinosis, avoiding side-effects such as anaphylaxis.
 SQ Sequence 17 AA:

Query Match 17.1% Score 138; DB 1; Length 17;
 Best Local Similarity 100.0%; Pred. No. 3,926-03;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 1 NRWLQFAKLIGFTLMG 17
 IIII IIII IIII IIII
 QY 55 NRWLQFAKLIGFTLMG 71

RESULT 4
 ID W80345 standard; peptide: 16 AA.
 AC W80345

DE 11-JAN-1999 (first entry)
 DE Sugi allergen protein CryJ2 derived epitope for T cells.
 KW T cell; epitope; sugi allergen; protein CryJ2; CryJ2; treatment;
 KW sugi; pollinosis; allergic reaction; pollen.
 OS Synthetic.
 PN J1035996-A.
 PD 29-SEP-1998.

PF 22-DEC-1997; 353448.
 PR 24-DEC-1996; JP-343441.
 RA (HAYB.) HAYASHIBARA SEIBUTSU KAGAKU.
 RI SANKYO CO. LTD.
 DR WP: 96-577037/49.
 PT A linked T cell epitope peptide - used for the treatment of
 PT Sugi-pollinosis
 PS Claim 7; Page 18; 21pp; Japanese.
 CC W80339-59 represent epitopes for T cells, derived from the sugi allergen
 CC proteins CryJ1 (W80339-46, W80350-53) and W80356-58) and CryJ2 (W80355-49
 CC and W80354-55). The peptides are useful for the treatment of
 CC Sugi-pollinosis. An allergic reaction of the body to follow.
 SQ Sequence 18 AA:

Query Match 17.0% Score 137; DB 1; Length 18;
 Best Local Similarity 94.1%; Pred. No. 4,026-03;
 Matches 16; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 1 RINWLQFAKLIGFTLMGK 17
 IIII IIII IIII
 QY 56 RINWLQFAKLIGFTLMGR 72

RESULT 5
 ID R81588 standard; Peptide: 17 AA.
 AC R81588
 DE 24-MAY-1996 (first entry)
 DE Cedar pollen allergen peptide 14 (homologue).
 KW Cedar; Cryptomeria japonica; pollen; allergen; immunoglobulin E;
 KW IgE; T-cell; epitope; antibody; pollinosis; therapy; immunotherapy.
 OS Synthetic.
 PN EP-700429-A2.
 PD 13-MAR-1996.
 PF 08-SEP-1995; 306295.
 PR 10-SEP-1994; JP-242137.
 RR 14-JUL-1995; JP-200221.
 RR 14-JUL-1995; JP-200204.
 RA (HAYB.) HAYASHIBARA SEIBUTSU KAGAKU.
 RI Hiko K., Saito S., Taniguchi Y.
 DR WP: 96-140976/15.
 PT New peptide(s) derived from cedar pollen allergens - activate
 PT allergen-specific T-cells, but not allergen-specific IgE antibodies.
 PI used for treating cedar pollinosis
 IS Disclosure; Page 32; 36pp; English.
 CC Synthetic peptides based on portions of cedar pollen allergens A
 CC (R81586) and B (R81587) were tested for their ability to activate
 CC cedar allergen-specific T-cells, but not allergen-specific IgE
 CC antibodies. 6 peptides (R81582-R81587) were identified as T-cell
 CC epitopes. These peptides, plus subsequences (R81588-95) essential
 CC for T-cell recognition, and homologous peptides (R81589-95) can
 CC be used as immunotherapeutic agents to treat or prevent cedar
 CC pollinosis, avoiding side-effects such as anaphylaxis.
 SQ Sequence 17 AA:

Query Match 16.3% Score 133; DB 1; Length 17;
 Best Local Similarity 94.1%; Pred. No. 1,266-02;
 Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 1 NRWLQFAKLIGFTLMG 17
 IIII IIII IIII
 QY 55 NRWLQFAKLIGFTLMG 71

RESULT 6
 ID W42171 standard; peptide: 20 AA.
 AC W42171
 DE 16-JUN-1998 (first entry)
 DE T-cell epitope peptide 51 from Japanese cypress pollen; antigen Chao2.
 KW Japanese cypress pollen; antigen; T-cell epitope; Chao1; Chao2;
 KW diagnosis; allergy; spring tree pollen disease; pollinosis.
 OS Cramaeyparis obtusa.
 PN W09747648-A1.

```

11 18-DEC-1997.
12 12-JUN-1997: J 2631.
13 14-JUN-1996: JP-151527.
14 (MEIP ) MEIJ MILK PROD CO LTD.
15 Daiichi K. Kinok.
16 WP1: 94-05242/05.
17 1-cell epitope peptide portion of Japanese cypress pollen antigens
18 Chao1 and Chao2 - used for diagnosis and treatment of spring tree
19 pollen disease.
20 Claim 2, Para 39: 71pp: Japanese.
21 The present sequence represents a T-cell epitope peptide from Japanese
22 cypress pollen antigen Chao2; the present invention describes peptides
23 which correspond to the T-cell epitope sites on Japanese cypress pollen
24 antigens Chao1 and Chao2. The peptides can be used as a reagent for the
25 diagnosis of allergy to Japanese cypress pollen, and as an antigen in
26 the treatment and prevention of spring tree pollen disease in which the
27 pollinosis involves reactivity to Japanese cypress pollen.
28 Sequence 20 AA:
29
30 Query Match 16.3% Score 134: DB 1: Length 20:
31 Best Local Similarity 73.7%: Pred. No. 1,25e-02:
32 Matches 14: Conservative 2: Mismatches 3: Indels 0: Gaps 0:
33
34 2 PWKSNKINQFAKLTGFLD 20
35 WP1: 94-05242/05
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193
2194
2195
2196
2197
2198
2199
2200
2201
2202
2203
2204
2205
2206
2207
2208
2209
2210
2211
2212
2213
2214
2215
2216
2217
2218
2219
2220
2221
2222
2223
2224
2225
2226
2227
2228
2229
2230
2231
2232
2233
2234
2235
2236
2237
2238
2239
2240
2241
2242
2243
2244
2245
2246
2247
2248
2249
2250
2251
2252
2253
2254
2255
2256
2257
2258
2259
2260
2261
2262
2263
2264
2265
2266
2267
2268
2269
2270
2271
2272
2273
2274
2275
2276
2277
2278
2279
2280
2281
2282
2283
2284
2285
2286
2287
2288
2289
2290
2291
2292
2293
2294
2295
2296
2297
2298
2299
2300
2301
2302
2303
2304
2305
2306
2307
2308
2309
2310
2311
2312
2313
2314
2315
2316
2317
2318
2319
2320
2321
2322
2323
2324
2325
2326
2327
2328
2329
2330
2331
2332
2333
2334
2335
2336
2337
2338
2339
2340
2341
2342
2343
2344
2345
2346
2347
2348
2349
2350
2351
2352
2353
2354
2355
2356
2357
2358
2359
2360
2361
2362
2363
2364
2365
2366
2367
2368
2369
2370
2371
2372
2373
2374
2375
2376
2377
2378
2379
2380
2381
2382
2383
2384
2385
2386
2387
2388
2389
2390
2391
2392
2393
2394
2395
2396
2397
2398
2399
2400
2401
2402
2403
2404
2405
2406
2407
2408
2409
2410
2411
2412
2413
2414
2415
2416
2417
2418
2419
2420
2421
2422
2423
2424
2425
2426
2427
2428
2429
2430
2431
2432
2433
2434
2435
2436
2437
2438
2439
2440
2441
2442
2443
2444
2445
2446
2447
2448
2449
2450
2451
2452
2453
2454
2455
2456
2457
2458
2459
2460
2461
2462
2463
2464
2465
2466
2467
2468
2469
2470
2471
2472
2473
2474
2475
2476
2477
2478
2479
2480
2481
2482
2483
2484
2485
2486
2487
2488
2489
2490
2491
2492
2493
2494
2495
2496
2497
2498
2499
2500
2501
2502
2503
2504
2505
2506
2507
2508
2509
2510
2511
2512
2513
2514
2515
2516
2517
2518
2519
2520
2521
2522
2523
2524
2525
2526
2527
2528
2529
2530
2531
2532
2533
2534
2535
2536
2537
2538
2539
2540
2541
2542
2543
2544
2545
2546
2547
2548
2549
2550
2551
2552
2553
2554
2555
2556
2557
2558
2559
2560
2561
2562
2563
2564
2565
2566
2567
2568
2569
2570
2571
2572
2573
2574
2575
2576
2577
2578
2579
2580
2581
2582
2583
2584
2585
2586
2587
2588
2589
2590
2591
2592
2593
2594
2595
2596
2597
2598
2599
2600
2601
2602
2
```



```

EP 08-SEP-1995: 306295.
FR 10-SEP-1994: JP-242137.
FR 14-JUL-1995: JP-200221.
FR 14-JUL-1995: JP-200234.
FA (HAYR : HAYASHIBARA SEIBUTSU KAGAKU.
PI Hino K., Saito S., Taniguchi Y.
CR WPI: 96-140476/15.
PI New peptide(s) derived from cedar pollen allergens - activate
PI allergen-specific T-cells, but not allergen-specific IgE antibodies.
PI used for treating cedar pollinosis.
PS Claim 4: Page 28: 36pp: English.
CC Synthetic peptides based on portions of cedar pollen allergens A
CC (R8156) and B (R8158) were tested for their ability to activate
CC cedar allergen-specific T-cells, but not allergen-specific IgE
CC antibodies. 6 peptides (R81580-R81585) were identified as T-cell
CC epitopes. These peptides, plus subsequence peptides (R81583-79), essential
CC for T-cell reconstitution, and homologous peptides (R81586-96) can
CC be used as immunotherapeutic agents to treat or prevent cedar
CC pollinosis, avoiding side-effects such as anaphylaxis.
SQ Sequence 17 AA:

Query Match 14.3% Score 115. DB 1: Length 17:
Residual Similarity 130.8% Pred. No: 2,51e-01:
Matches 15: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

EP 1 EFASKNPHLOKNTDIT 17
    IIII IIIII
QY 32 EFASKNPHLOKNTDIT 48

RESULT 15
EP 08-SEP-1995: 306295.
FR 10-SEP-1994: JP-242137.
FR 14-JUL-1995: JP-200221.
FR 14-JUL-1995: JP-200234.
FA (HAYR : HAYASHIBARA SEIBUTSU KAGAKU
PI Hino K., Saito S., Taniguchi Y.
CR WPI: 96-140476/15.
PI New peptide(s) derived from cedar pollen allergens - activate
PI allergen-specific T-cells, but not allergen-specific IgE antibodies.
PI used for treating cedar pollinosis.
PS Claim 4: Page 28: 36pp: English.
CC Synthetic peptides based on portions of cedar pollen allergens A
CC (R8156) and B (R8158) were tested for their ability to activate
CC cedar allergen-specific T-cells, but not allergen-specific IgE
CC antibodies. 6 peptides (R81580-R81585) were identified as T-cell
CC epitopes. These peptides, plus subsequence peptides (R81583-79), essential
CC for T-cell reconstitution, and homologous peptides (R81586-96) can
CC be used as immunotherapeutic agents to treat or prevent cedar
CC pollinosis, avoiding side-effects such as anaphylaxis.
SQ Sequence 17 AA:

Query Match 13.8% Score 111. DB 1: Length 17:
Residual Similarity 93.8% Pred. No: 5,22e-01:
Matches 15: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

Dr 1 KVDGIIAAVNPASWK 16
    I IIII IIIII
QY 40 KVDGIIAAVNPASWK 105

```

.....
[W][E][B][S][I][T][E]
..... (TM)
.....

Release 3.1A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPsrch_pp protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Jun 20 13:38:29 2000: MaxPar time 4.18 Seconds
352.971 Million cell updates/sec

Tabular output not generated.

Title: US-09-142-524A-2
Description: (1-105) from US09142524A.pep
Perfect Score: 808
Sequence: 1 MKVTVAENQFGPNRVEIKR.....FQGRVDGIIAAYQNFASNK 105

Scoring table: PAM 150
Gap 11

Searched: 145341 seqs, 14437480 residues

Post-processing: Minimum Match 08
Listing first 45 summaries
Maximum DB seq length 105

Databases: 3-issued
1:5A_COMB 2:5B_COMB 3:4_COMB 4:PGT_COMB 5:backfiles1

Statistics: Mean 27.828; Variance 123.399; scale 0.226

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

*
SUMMARIES

| Result
No. | Score | Query
Match | Length | DB | ID | Description | Prod. No. |
|---------------|-------|----------------|--------|----|----|-------------|-----------|
|---------------|-------|----------------|--------|----|----|-------------|-----------|

.....

No matches found.

Search completed: Tue Jun 20 13:38:33 2000
Job time : 5 secs.



Release 3.0A John F. Collins, BioComputing Research Unit.
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Mitsch_PP protein - protein database search, using Smith-Waterman algorithm
Run on: Mon Jun 19 15:59:18 2000 Master time 9.36 Seconds
529,224 Million cell updates/sec
Tutular output not generated.
Title: US-09-142-524A-2
Description: (1-105) from US9142524A.ppt
Perfect Score: 806
Sequence: 1 MKVTVAFNQGPNRVRFKR.....FDGRVDGIIAAYQNPSWK 105
Scoring table: FAM 100
Gap 11
Searched: 54280 seqs, 47172406 residues
Post-processing: Minimum Match 0%
Listing first 45 summaries
Maximum FP seq length 105
Database: par62
pir1 2:pir2 3:pir3 4:pir4
Statistics: Mean 41.762 Variance 84.448 scale 0.495
Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.
Result Quality
No. Score Match Length ID Description Prod. No.

SUMMARIES

No matches found.
Search completed: Mon Jun 19 15:59:36 2000
Job time : 18 secs.

(五)

20

protein - protein database search, using Smith-Waterman algorithm

Run on: 19 15:57:59 Maspar time 6.72 seconds

| Run On: | Run Off: | WASPd Time | WASPd Cell Updates/sec |
|----------|----------|------------|------------------------|
| 13:57:35 | 13:57:55 | 2000 | 475.570 |

Tubular output not generated.

File: 2JS-09-142-524A-2

Description: (1-105) FROM US09142524A.pep

Verdicts and

Sequence: 1 MKVTVAENCFGNRRVF¹KR.....FDGRVCGIIAAYQNFA¹⁰⁵SWK

Source: L.A. 100 PAM: 50

| Case | Year |
|------|------|
| 1 | 1990 |
| 2 | 1991 |
| 3 | 1992 |
| 4 | 1993 |
| 5 | 1994 |
| 6 | 1995 |
| 7 | 1996 |
| 8 | 1997 |
| 9 | 1998 |
| 10 | 1999 |
| 11 | 2000 |
| 12 | 2001 |
| 13 | 2002 |
| 14 | 2003 |
| 15 | 2004 |
| 16 | 2005 |
| 17 | 2006 |
| 18 | 2007 |
| 19 | 2008 |
| 20 | 2009 |
| 21 | 2010 |
| 22 | 2011 |
| 23 | 2012 |
| 24 | 2013 |
| 25 | 2014 |
| 26 | 2015 |
| 27 | 2016 |
| 28 | 2017 |
| 29 | 2018 |
| 30 | 2019 |
| 31 | 2020 |
| 32 | 2021 |
| 33 | 2022 |
| 34 | 2023 |
| 35 | 2024 |
| 36 | 2025 |
| 37 | 2026 |
| 38 | 2027 |
| 39 | 2028 |
| 40 | 2029 |
| 41 | 2030 |
| 42 | 2031 |
| 43 | 2032 |
| 44 | 2033 |
| 45 | 2034 |
| 46 | 2035 |
| 47 | 2036 |
| 48 | 2037 |
| 49 | 2038 |
| 50 | 2039 |
| 51 | 2040 |
| 52 | 2041 |
| 53 | 2042 |
| 54 | 2043 |
| 55 | 2044 |
| 56 | 2045 |
| 57 | 2046 |
| 58 | 2047 |
| 59 | 2048 |
| 60 | 2049 |
| 61 | 2050 |
| 62 | 2051 |
| 63 | 2052 |
| 64 | 2053 |
| 65 | 2054 |
| 66 | 2055 |
| 67 | 2056 |
| 68 | 2057 |
| 69 | 2058 |
| 70 | 2059 |
| 71 | 2060 |
| 72 | 2061 |
| 73 | 2062 |
| 74 | 2063 |
| 75 | 2064 |
| 76 | 2065 |
| 77 | 2066 |
| 78 | 2067 |
| 79 | 2068 |
| 80 | 2069 |
| 81 | 2070 |
| 82 | 2071 |
| 83 | 2072 |
| 84 | 2073 |
| 85 | 2074 |
| 86 | 2075 |
| 87 | 2076 |
| 88 | 2077 |
| 89 | 2078 |
| 90 | 2079 |
| 91 | 2080 |
| 92 | 2081 |
| 93 | 2082 |
| 94 | 2083 |
| 95 | 2084 |
| 96 | 2085 |
| 97 | 2086 |
| 98 | 2087 |
| 99 | 2088 |
| 100 | 2089 |
| 101 | 2090 |
| 102 | 2091 |
| 103 | 2092 |
| 104 | 2093 |
| 105 | 2094 |
| 106 | 2095 |
| 107 | 2096 |
| 108 | 2097 |
| 109 | 2098 |
| 110 | 2099 |
| 111 | 2100 |
| 112 | 2101 |
| 113 | 2102 |
| 114 | 2103 |
| 115 | 2104 |
| 116 | 2105 |
| 117 | 2106 |
| 118 | 2107 |
| 119 | 2108 |
| 120 | 2109 |
| 121 | 2110 |
| 122 | 2111 |
| 123 | 2112 |
| 124 | 2113 |
| 125 | 2114 |
| 126 | 2115 |
| 127 | 2116 |
| 128 | 2117 |
| 129 | 2118 |
| 130 | 2119 |
| 131 | 2120 |
| 132 | 2121 |
| 133 | 2122 |
| 134 | 2123 |
| 135 | 2124 |
| 136 | 2125 |
| 137 | 2126 |
| 138 | 2127 |
| 139 | 2128 |
| 140 | 2129 |
| 141 | 2130 |
| 142 | 2131 |
| 143 | 2132 |
| 144 | 2133 |
| 145 | 2134 |
| 146 | 2135 |
| 147 | 2136 |
| 148 | 2137 |
| 149 | 2138 |
| 150 | 2139 |
| 151 | 2140 |
| 152 | 2141 |
| 153 | 2142 |
| 154 | 2143 |
| 155 | 2144 |
| 156 | 2145 |
| 157 | 2146 |
| 158 | 2147 |
| 159 | 2148 |
| 160 | 2149 |
| 16 | |

Searched: 83957 seqs, 30454973 residues

Post-processing: Minimum Match 98

Listing : 1155 45 summaries

Maximum DE seq length 105

Database: swiss-prot:38

101 DISTRICT

Statistics. Mean 42.649; Variance 75.460; scale 0.565

Pred. No. is the number of results predicted, chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result | Score | Grady
Match | Location | Date | Prod. No. |
|--------|-------|----------------|----------|------|-----------|
| | | | | | |

No matches found.

search completed: Mon Jun :9 15:58:19 2000

Job time : 21 sec.

.....
[W][E][B][S][I][T][E]
..... (TM)
.....

Release 3.1A John F. Collins, BioComputing Research Unit,
Copyright (c) 1991-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Wsearch.ppp Protein - protein database search, using Smith-Waterman algorithm
Run on: Mon Jun 19 15:58:35 2000 MasPar time 15.33 Seconds
474.840 Million cell updates/sec
Tabular output not generated.

Title: US-09-142-524A-2
Description: (1:105) from USC9:42524A.ppp
Perfect Score: 806
Sequence: 1 MKVTVAENGFGPNRRVFIRK.....FQGRVGGIIAAYQNPASWK 105

Scoring table: PAM 150
Gap 11

Searched: 225879 seqs, 69334122 residues
Post-processing: Minimum Match 08
Listing first 45 summaries
Maximum DB seq length 105

Database: spiremb12
1:sp_archaea 2:sp_bacteria 3:sp_fungal 4:sp_human
5:sp_invertebrate 6:sp_mammal 7:sp_mhc 8:sp_organelle
9:sp_phase 10:sp_plant 11:sp_rodent 12:sp_unclassified
13:sp_vertebrate 14:sp_virus

Statistics: Mean 41.678; Variance 72.759; Scale 0.573

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result | Query | No. | Score | Match | Length | ID | Description | Pred. No. |
|--------|-------|-----|-------|-------|--------|----|-------------|-----------|
|--------|-------|-----|-------|-------|--------|----|-------------|-----------|

.....

No matches found.

Search completed: Mon Jun 19 15:59:02 2000
Job time : 27 secs.

.....

WATERMANN

(TM)

.....

Release 1.1A John F. Collins, BioComputing Research Unit,
Copyright (c) 1994-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Search: protein protein database search, using Smith-Waterman algorithm
Run On: Mon Jun 19 18:58:37 2000: MasPar time 6.77 Seconds
468,570 Million cell updates/sec
format output not deprecated.

File: >US-09-142-524A-3
Description: (2134), from US09142524A.pep
RefSeq Score: 1026
Sequence: 1 MKVTVAQFGNNRVFKR.....KFIKRVCGIIAAYONPASWK 134

Scoring table: PAM 150
Gap 11

Search: 188,663 seqs, 23686126 residues

Post-processing: Minimum Match 08
Listing first 45 summaries

Database: database35
Language: English

Statistics: Mean 30.69% Variance 132.26% scale 0.232

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description | Pred. No. |
|------------|-------|-------------|--------|----------|-------------------------|-----------|
| 1 | 1026 | 100.0 | 134 | 1 W27371 | Multi-epitope peptide | 4.34e-87 |
| 2 | 961 | 94.7 | 105 | 1 W27372 | Multi-epitope peptide | 6.58e-42 |
| 3 | 772 | 66.3 | 80 | 1 W27369 | Multi-epitope peptide | 5.64e-24 |
| 4 | 493 | 26.3 | 353 | 1 R51388 | Japanese cedar pollen | 6.30e-17 |
| 5 | 296 | 28.8 | 353 | 1 R51587 | Cedar pollen allergen | 6.30e-17 |
| 6 | 296 | 28.8 | 374 | 1 R41937 | Cry j 1 | 6.30e-17 |
| 7 | 296 | 28.8 | 374 | 1 R61665 | Japanese cedar pollen | 6.30e-17 |
| 8 | 296 | 28.8 | 374 | 1 R45541 | Cry j 1 pollen allergen | 6.30e-17 |
| 9 | 296 | 28.8 | 374 | 1 R2490 | Cry j 1 Japanese Cedar | 6.30e-17 |
| 10 | 224 | 21.8 | 367 | 1 R45572 | Jun s 1 | 2.07e-10 |
| 11 | 224 | 21.8 | 370 | 1 R45578 | Jun v 1 | 2.07e-10 |
| 12 | 210 | 20.5 | 354 | 1 W2121 | Japanese cypress pollen | 3.61e-09 |
| 13 | 210 | 20.5 | 354 | 1 W4344 | Chamaecyparis obtusa p | 3.61e-09 |
| 14 | 210 | 20.5 | 375 | 1 W4345 | Chamaecyparis obtusa p | 5.42e-09 |
| 15 | 208 | 20.3 | 30 | 1 W4682 | T-cell epitope peptide | 5.42e-09 |
| 16 | 172 | 16.8 | 460 | 1 R69731 | Japonicum allergenes | 7.32e-06 |
| 17 | 172 | 16.8 | 514 | 1 R21585 | Cedar pollen allergen | 7.32e-06 |
| 18 | 172 | 16.8 | 514 | 1 R31690 | Japanese cedar pollen | 7.32e-06 |
| 19 | 172 | 16.8 | 514 | 1 R31599 | Japan cedar pollen Cry | 7.32e-06 |
| 20 | 172 | 16.8 | 514 | 1 R74333 | Japanese cedar pollen | 7.32e-06 |
| 21 | 172 | 16.8 | 514 | 1 R59792 | Japonicum allergen | 7.32e-06 |
| 22 | 168 | 15.4 | 20 | 1 R45549 | Cry j 1 pollen allergen | 1.13e-04 |
| 23 | 168 | 15.4 | 20 | 1 R2498 | Cry j 1 Japanese Cedar | 1.13e-04 |

| | | | | | | |
|----|-----|------|-----|----------|----------------------------|----------|
| 24 | 149 | 14.5 | 514 | 1 W43446 | Chamaecyparis obtusa p | 5.45e-04 |
| 25 | 149 | 14.5 | 514 | 1 W42122 | Japanese cypress pollen | 5.45e-04 |
| 26 | 144 | 14.0 | 397 | 1 R43565 | Ragweed Pollen Allergen | 1.48e-03 |
| 27 | 142 | 13.8 | 30 | 1 W45579 | Cry j 1 pollen allergen | 2.45e-03 |
| 28 | 141 | 13.7 | 387 | 1 W45851 | Ragweed pollen Amb a 1 | 2.97e-03 |
| 29 | 141 | 13.7 | 387 | 1 W4333 | Amb a 1/Antigen E | 2.97e-03 |
| 30 | 141 | 13.7 | 387 | 1 W4727 | Allergen Amb a 1A | 2.97e-03 |
| 31 | 141 | 13.7 | 396 | 1 W2505 | Ragweed Amb a 1 allergen | 2.97e-03 |
| 32 | 141 | 13.7 | 396 | 1 W4564 | Ragweed Pollen Allergen | 2.97e-03 |
| 33 | 141 | 13.7 | 397 | 1 W4245 | Amb a 1/Antigen E | 2.97e-03 |
| 34 | 141 | 13.7 | 397 | 1 W58729 | Allergen Amb a 1C | 2.97e-03 |
| 35 | 141 | 13.7 | 397 | 1 W62507 | Ragweed Amb a 1.3 allergen | 2.97e-03 |
| 36 | 141 | 13.7 | 431 | 1 W41653 | Ragweed pollen Ant a 1 | 2.97e-03 |
| 37 | 138 | 13.5 | 17 | 1 R81581 | Cedar pollen allergen | 5.24e-03 |
| 38 | 139 | 13.5 | 294 | 1 R07328 | Part of Amb a 1/Antigen | 4.34e-03 |
| 39 | 137 | 13.4 | 18 | 1 W80345 | Supl allergen protein | 6.46e-03 |
| 40 | 137 | 13.4 | 391 | 1 R07327 | Part of Amb a 1/Antigen | 6.46e-03 |
| 41 | 137 | 13.4 | 398 | 1 W58728 | Allergen Amb a 1B | 6.46e-03 |
| 42 | 137 | 13.4 | 398 | 1 R45545 | Ragweed Pollen Allergen | 6.46e-03 |
| 43 | 137 | 13.4 | 398 | 1 R07334 | Amb a 1/Antigen E | 6.46e-03 |
| 44 | 137 | 13.4 | 398 | 1 W2506 | Ragweed Amb a 1.2 allergen | 6.46e-03 |
| 45 | 137 | 13.4 | 442 | 1 W1855 | Ragweed pollen GNG class | 5.33e-03 |

ALIGNMENTS

RESULT 1
ID W27371 standard: peptide: 134 AA.
AC W27371
DI 24-MAR-1998 (first entry)
DE Multi-epitope peptide used as immunotherapeutic agent #1
KW Multi-epitope peptide; immunotherapeutic agent; allergic disease;
KW T-cell epitope region; allergen; lymphocyte; immunoglobulin E;
PS Synthetic
OS W09732600.A1
PD 12-SEP-1997
PF 10-MAR-1997: JC0740
PR 10-MAR-1996: JP-080702
PA (MEIP) MEICI MILK PROD CO LTD,
PI Dairiki K, Iwara A, Kino K, Kame A, Sone T
DR WPI: 97-470495/43
PR Peptide immunotherapeutic agent to treat allergic diseases
PR Contains multi-epitope peptide containing T cell epitope regions
PR from different allergens
PS Claim 5: Page 32: 58pp: Japanese
CC The present sequence represents a multi-epitope peptide which is used as
CC a new immunotherapeutic agent. It comprises 1 cell epitope regions from 2
CC or more different allergens (preferably linked via aromatic lysine
CC dimers), where the T cell epitope regions have a positivity index
CC greater than 100 as measured in a patient group responding to the
CC allergen. have at least 70% reactivity with lymphocytes from patients
CC responding to the allergen; and are not reactive with immunoglobulin E
CC (IgE) antibodies from patients responsive to the allergen. The agent can
CC be used to prevent and treat a wide variety of allergic diseases, e.g. by
CC desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
SQ Sequence 134 AA:
Query Match 100.0% Score 1026; DB 1; Length 134;
Best Local Similarity 100.0% Pred. No. 4.34e-87;
Matches 134; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | |
|----|-----|---|-----|
| Db | 1 | MKVTVAFNQFGNNRVFKRVSNNVHGRIRIDIFASKNHFLOKNTIGTRKNNRWLQ | 60 |
| QY | 1 | MKVTVAFNQFGNNRVFKRVSNNVHGRIRIDIFASKNHFLOKNTIGTRKNNRWLQ | 60 |
| Db | 61 | FAKLGTGLMGRPRKLIIFSGNNMKLKMPLYIAGYKTFDGRRAEVSVYHVGAKFIRRV | 120 |
| QY | 61 | FAKLGTGLMGRPRKLIIFSGNNMKLKMPLYIAGYKTFDGRRAEVSVYHVGAKFIRRV | 120 |
| Db | 121 | DCI:IAAYONPASWK 134 | |
| QY | 121 | DCI:IAAYONPASWK 134 | |

HA (HAYB) HAYASHIPARA SEIBUSU KAGAKU.
 HI Hino K. Saito S. Taniguchi Y.
 WI: 96-140575/15.

PI New peptide(s) derived from cedar pollen allergens - activate
 PI allergen specific T-cells, but not allergen-specific IgE antibodies.
 PI used for treating cedar pollinosis.

PS Claim 5: Para 432: 4pp: English.
 SN Synthetic peptides based on portions of cedar pollen allergens A
 SN (R81586) and B (R81587) were tested for their ability to activate
 SN cedar allergen-specific T-cells, but not allergen-specific IgE
 SN antibodies. 2 peptides (R81580-R81585) were identified as T-cell
 SN epitopes. These peptides, plus subsequences (R81571-79) essential
 SN for T-cell recognition and homologous peptides (R81588-96) can
 SN be used as immunotherapeutic agents to treat or prevent cedar
 SN pollinosis, avoiding side-effects such as anaphylaxis.
 SN Sequence 374 AA:

Query Match 28.8% Score 296: DB 1: Length 374:

Best Local Similarity 73.6% Pred. No. 6.30e-17:

Matches 39: Conservative 5: Mismatches 7: Indels 2: Gaps 2:

DB 66 RPLWIFSGNNK:KLKMPYIAGYKTFDGRGAQV-YIGNGPCVFIKRVSNV: 117

WI: 96-140575/15

DB 73 RPLWIFSGNNK:KLKMPYIAGYKTFDGRGAQV-YIGNGPCVFIKRVSNV: 124

Query Match 28.8% Score 296: DB 1: Length 374:

Best Local Similarity 73.6% Pred. No. 6.30e-17:

Matches 39: Conservative 5: Mismatches 7: Indels 2: Gaps 2:

DB 66 RPLWIFSGNNK:KLKMPYIAGYKTFDGRGAQV-YIGNGPCVFIKRVSNV: 117

WI: 96-140575/15

DB 73 RPLWIFSGNNK:KLKMPYIAGYKTFDGRGAQV-YIGNGPCVFIKRVSNV: 124

Query Match 28.8% Score 296: DB 1: Length 374:

Best Local Similarity 73.6% Pred. No. 6.30e-17:

Matches 39: Conservative 5: Mismatches 7: Indels 2: Gaps 2:

DB 66 RPLWIFSGNNK:KLKMPYIAGYKTFDGRGAQV-YIGNGPCVFIKRVSNV: 117

WI: 96-140575/15

DB 73 RPLWIFSGNNK:KLKMPYIAGYKTFDGRGAQV-YIGNGPCVFIKRVSNV: 124

Query Match 28.8% Score 296: DB 1: Length 374:

Best Local Similarity 73.6% Pred. No. 6.30e-17:

Matches 39: Conservative 5: Mismatches 7: Indels 2: Gaps 2:

DB 66 RPLWIFSGNNK:KLKMPYIAGYKTFDGRGAQV-YIGNGPCVFIKRVSNV: 117

WI: 96-140575/15

DB 73 RPLWIFSGNNK:KLKMPYIAGYKTFDGRGAQV-YIGNGPCVFIKRVSNV: 124

Query Match 28.8% Score 296: DB 1: Length 374:

Best Local Similarity 73.6% Pred. No. 6.30e-17:

Matches 39: Conservative 5: Mismatches 7: Indels 2: Gaps 2:

DB 66 RPLWIFSGNNK:KLKMPYIAGYKTFDGRGAQV-YIGNGPCVFIKRVSNV: 117

WI: 96-140575/15

DB 73 RPLWIFSGNNK:KLKMPYIAGYKTFDGRGAQV-YIGNGPCVFIKRVSNV: 124

Query Match 28.8% Score 296: DB 1: Length 374:

Best Local Similarity 73.6% Pred. No. 6.30e-17:

Matches 39: Conservative 5: Mismatches 7: Indels 2: Gaps 2:

DB 66 RPLWIFSGNNK:KLKMPYIAGYKTFDGRGAQV-YIGNGPCVFIKRVSNV: 117

WI: 96-140575/15

DB 73 RPLWIFSGNNK:KLKMPYIAGYKTFDGRGAQV-YIGNGPCVFIKRVSNV: 124

Query Match 28.8% Score 296: DB 1: Length 374:

Best Local Similarity 73.6% Pred. No. 6.30e-17:

Matches 39: Conservative 5: Mismatches 7: Indels 2: Gaps 2:

DB 66 RPLWIFSGNNK:KLKMPYIAGYKTFDGRGAQV-YIGNGPCVFIKRVSNV: 117

WI: 96-140575/15

DB 73 RPLWIFSGNNK:KLKMPYIAGYKTFDGRGAQV-YIGNGPCVFIKRVSNV: 124

Query Match 28.8% Score 296: DB 1: Length 374:

Best Local Similarity 73.6% Pred. No. 6.30e-17:

Matches 39: Conservative 5: Mismatches 7: Indels 2: Gaps 2:

DB 66 RPLWIFSGNNK:KLKMPYIAGYKTFDGRGAQV-YIGNGPCVFIKRVSNV: 117

WI: 96-140575/15

DB 73 RPLWIFSGNNK:KLKMPYIAGYKTFDGRGAQV-YIGNGPCVFIKRVSNV: 124

DE Japanese cedar pollen antigen Cry j I.
 DE Japanese cedar pollen antigen: allergen: Cry j I: sugi pollenosis
 OS Cryptomeria japonica.
 FH Key Location/Qualifiers
 FI peptide 1: 21
 FI protein 22: 374
 FI /label= signal_peptide
 FI /label= mature_Cryj
 FI J06197768-A.
 PN 19-JUL-1994.
 PF 07-JAN-1993: C01116
 PR 07-JAN-1993: JP-001116.
 PA (MEL) MEJJI SEIKA KAISHA.
 DR WP1: 94-268680/33.
 DR N-PSDB: Q71601.
 PT Sugi (Japanese cedar) pollen antigen Cry j I: as useful for
 PT diagnosis, treatment and prevention of sugi pollenosis
 PS Claim 2: Page 5-7: 9pp: Japanese.
 CC The coding sequence for the Japanese cedar ("sugi") pollen allergen
 CC Cry j I was isolated from a cDNA library prepared from POLY A+ RNA. All
 CC or part of the Cry j I protein can be used for diagnosis, treatment
 CC and prevention of sugi pollenosis.
 SQ Sequence 374 AA:

Query Match 28.8% Score 296: DB 1: Length 374:

Best Local Similarity 73.6% Pred. No. 6.30e-17:

Matches 39: Conservative 5: Mismatches 7: Indels 2: Gaps 2:

DB 87 RPLWIFSGNNK:KLKMPYIAGYKTFDGRGAQV-YIGNGPCVFIKRVSNV: 118

WI: 96-140575/15

DB 73 RPLWIFSGNNK:KLKMPYIAGYKTFDGRGAQV-YIGNGPCVFIKRVSNV: 124

RESULT 8

ID R45541 standard: Protein: 374 AA.

AC R45541:

DE 13-JUL-1994 (first entry)

DE Cry j I pollen allergen.

KW Japanese cedar: detection: allergy: treatment: diagnosis:

KW T cell epitope: sensitivity:

OS Cryptomeria japonica.

FH Key Location/Qualifiers

FI peptide 1: 21

FI /note= "signal peptide"

FI peptide 22: 374

FI /note= "mature peptide"

PN W09401560-A.

PD 20-JAN-1994.

PF 15-JAN-1993: C01116

PR 10-JUL-1992: W0-05561.

PR 07-SEP-1992: US-938960.

PA (IMMO-) IMMLOGIC PHARM CORP.

PI Bend JF. Garman RD. Griffith JJ. Kuo M. Pollock J.

DR WP1: 94-035066/04.

DR N-PSDB: Q5271.

PT Antigens derived from Japanese cedar pollen allergen Cry j I.

PT contain at least two T cell epitope(s). used to treat or diagnose

PT allergy

PS Disclosure: Fig 4: 137pp: English

CC The sequence is that of the Japanese cedar pollen allergen.

CC Cry j I which contains at least two T cell epitopes. Peptide

CC antigens derived from it can be used for the treatment and

CC diagnosis of allergies associated with Japanese cedar pollen.

CC The peptides have enhanced therapeutic properties but reduced

CC side effects compared to naturally occurring allergens.

SC Sequence 374 AA:

Query Match 28.8% Score 296: DB 1: Length 374:

Best Local Similarity 73.6% Pred. No. 6.30e-17:

Matches 39: Conservative 5: Mismatches 7: Indels 2: Gaps 2:

DB 87 RPLWIFSGNNK:KLKMPYIAGYKTFDGRGAQV-YIGNGPCVFIKRVSNV: 118

WI: 96-140575/15

DB 73 RPLWIFSGNNK:KLKMPYIAGYKTFDGRGAQV-YIGNGPCVFIKRVSNV: 124

Example 1: Page 14-15: 71pp: Japanese.

MORPHOLIN

(TM)

MORPHOLIN protein - protein database search, using Smith-Waterman algorithm
Run on: Tue Jun 20 12:33:35 2000: MasPar time 4.72 Seconds
410.196 Million cell updates/sec
Tabular output not generated.

Title: >US-09-142-524A-3
Description: (134) from US09142524A.ppt
Perfect Score: 1026
Sequence: 1 MKVTAFNPQPNRRVFIK.....KFIRVIGIYAAYNPASWK 134

Scoring table: PAM 150
Gap 11

Searched: 14541 seqs, 14437480 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: a-issued
1:5A_COMB 2:5B_COMB 3:6_COMB 4:PCT_COMB 5:backfiles1

Statistics: Mean 28.726; Variance 129.33; scale 0.222

Pred No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description | Pred. No. |
|------------|-------|-------------|--------|-------|-----------------------------------|-----------|
| 1 | 141 | 13.7 | 397 | 1 | US-08-175- Sequence 72, Applicat | 2.06e-03 |
| 2 | 141 | 13.7 | 387 | 1 | US-08-230- Sequence 72, Applicat | 2.06e-03 |
| 3 | 141 | 13.7 | 387 | 1 | US-08-250- Sequence 72, Applicat | 2.06e-03 |
| 4 | 141 | 13.7 | 397 | 1 | US-08-175- Sequence 76, Applicat | 2.06e-03 |
| 5 | 141 | 13.7 | 397 | 1 | US-08-230- Sequence 76, Applicat | 2.06e-03 |
| 6 | 141 | 13.7 | 397 | 1 | US-08-250- Sequence 76, Applicat | 2.06e-03 |
| 7 | 137 | 13.4 | 391 | 1 | US-08-175- Sequence 59, Applicat | 4.31e-03 |
| 8 | 137 | 13.4 | 391 | 1 | US-08-230- Sequence 59, Applicat | 4.31e-03 |
| 9 | 137 | 13.4 | 391 | 1 | US-08-250- Sequence 59, Applicat | 4.31e-03 |
| 10 | 137 | 13.4 | 398 | 1 | US-08-175- Sequence 74, Applicat | 4.31e-03 |
| 11 | 137 | 13.4 | 398 | 1 | US-08-230- Sequence 74, Applicat | 4.31e-03 |
| 12 | 137 | 13.4 | 398 | 1 | US-08-250- Sequence 74, Applicat | 4.31e-03 |
| 13 | 136 | 13.3 | 383 | 1 | US-08-175- Sequence 78, Applicat | 5.18e-03 |
| 14 | 136 | 13.3 | 383 | 1 | US-08-230- Sequence 78, Applicat | 5.18e-03 |
| 15 | 136 | 13.3 | 383 | 1 | US-08-250- Sequence 78, Applicat | 5.18e-03 |
| 16 | 124 | 12.1 | 383 | 1 | US-08-175- Sequence 80, Applicat | 4.62e-02 |
| 17 | 124 | 12.1 | 383 | 1 | US-08-230- Sequence 80, Applicat | 4.62e-02 |
| 18 | 124 | 12.1 | 383 | 1 | US-08-250- Sequence 80, Applicat | 4.62e-02 |
| 19 | 90 | 8.8 | 1079 | 3 | US-08-353- Sequence 8, Applicatio | 1.71e-01 |
| 20 | 90 | 8.8 | 1079 | 3 | US-08-485- Sequence 8, Applicatio | 1.71e-01 |
| 21 | 90 | 8.8 | 1079 | 2 | US-08-943- Sequence 8, Applicatio | 1.71e-01 |
| 22 | 90 | 8.8 | 1079 | 2 | US-08-480- Sequence 8, Applicatio | 1.71e-01 |
| 23 | 90 | 8.8 | 1079 | 1 | US-08-484- Sequence 8, Applicatio | 1.71e-01 |

| | | | | | | |
|----|----|-----|------|---|-----------------------------------|----------|
| 24 | 90 | 8.8 | 1085 | 3 | US-08-353- Sequence 5, Applicatio | 1.71e-01 |
| 25 | 90 | 8.8 | 1085 | 1 | US-08-485- Sequence 5, Applicatio | 1.71e-01 |
| 26 | 90 | 8.8 | 1085 | 2 | US-08-943- Sequence 5, Applicatio | 1.71e-01 |
| 27 | 90 | 8.8 | 1085 | 2 | US-08-485- Sequence 5, Applicatio | 1.71e-01 |
| 28 | 90 | 8.8 | 1085 | 1 | US-08-484- Sequence 5, Applicatio | 1.71e-01 |
| 29 | 89 | 8.7 | 1078 | 2 | US-08-484- Sequence 7, Applicatio | 2.02e-01 |
| 30 | 89 | 8.7 | 1078 | 2 | US-08-485- Sequence 7, Applicatio | 2.02e-01 |
| 31 | 89 | 8.7 | 1078 | 3 | US-08-353- Sequence 7, Applicatio | 2.02e-01 |
| 32 | 89 | 8.7 | 1078 | 1 | US-08-485- Sequence 7, Applicatio | 2.02e-01 |
| 33 | 89 | 8.7 | 1078 | 2 | US-08-943- Sequence 7, Applicatio | 2.02e-01 |
| 34 | 89 | 8.7 | 1088 | 3 | US-08-353- Sequence 6, Applicatio | 2.02e-01 |
| 35 | 89 | 8.7 | 1088 | 2 | US-08-943- Sequence 6, Applicatio | 2.02e-01 |
| 36 | 89 | 8.7 | 1088 | 1 | US-08-485- Sequence 6, Applicatio | 2.02e-01 |
| 37 | 89 | 8.7 | 1088 | 1 | US-08-484- Sequence 6, Applicatio | 2.02e-01 |
| 38 | 89 | 8.7 | 1088 | 2 | US-08-480- Sequence 6, Applicatio | 2.02e-01 |
| 39 | 89 | 8.7 | 1219 | 2 | US-08-687- Sequence 6, Applicatio | 2.02e-01 |
| 40 | 86 | 8.4 | 490 | 4 | PCT-US95-0 Sequence 7, Applicatio | 3.30e-01 |
| 41 | 86 | 8.4 | 490 | 1 | US-08-201- Sequence 7, Applicatio | 3.30e-01 |
| 42 | 86 | 8.4 | 490 | 2 | US-08-238- Sequence 7, Applicatio | 3.30e-01 |
| 43 | 86 | 8.4 | 802 | 2 | US-08-326- Sequence 4, Applicatio | 3.30e-01 |
| 44 | 86 | 8.4 | 802 | 1 | US-08-465- Sequence 5, Applicatio | 3.30e-01 |
| 45 | 86 | 8.4 | 802 | 1 | US-08-261- Sequence 2, Applicatio | 3.30e-01 |

ALIGNMENTS

RESULT 1
XX US-08-175-069A-72 STANDARD: PRT: 367 AA.
AC xxxxxx
DT
XX
DE
XX
XX Sequence 72, Application US/08175069A
CC Sequence 72, Application US/08175069A
CC Patent No. 577676:
CC GENERAL INFORMATION:
CC APPLICANT: Rogers, Bruce
CC APPLICANT: Klapper, David G.
CC APPLICANT: Rafnar, Thorunn
CC APPLICANT: Kuo, Wei-Chang
CC TITLE OF INVENTION: Allergenic Proteins From Recombinant FISH
CC NUMBER OF SEQUENCES: 9
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: LAHVE, S. J. KATHOLIC CH
CC STREET: 60 STATE STREET
CC CITY: Boston
CC STATE: Massachusetts
CC COUNTRY: USA
CC ZIP: 02109-1875
CC COMPUTER READABLE FORM: disk
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/175,069A
CC FILING DATE: December 29, 1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/529,951
CC FILING DATE: May 29, 1990
CC APPLICATION NUMBER: US 07/325,365
CC FILING DATE: March 17, 1989
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Amy E. Mandragouras
CC REGISTRATION NUMBER: 36,207
CC REFERENCE/DOCKET NUMBER: IMI-0180V
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617)227-7400
CC TELEFAX: (617)227-5941
CC INFORMATION FOR SEQ ID NO: 72:
CC SEQUENCE CHARACTERISTICS:

DE - Sequence 76, Application US/08175069A
XX
CC Sequence 76, Application US/08175069A
CC Patent No. 5776761
CC
CC GENERAL INFORMATION:
CC APPLICANT: Rogers, Bruce
CC APPLICANT: Klapper, David G.
CC APPLICANT: Rafnar, Thorunn
CC APPLICANT: Kuo, Mei-chang
CC TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses
CC NUMBER OF SEQUENCES: 93
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: LAHIVE & COCKFIELD, LLP
CC STREET: 60 State Street
CC CITY: Boston
CC STATE: Massachusetts
CC COUNTRY: USA
CC ZIP: 02109-1875
CC
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent in Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/175-069A
CC FILING DATE: December 29, 1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/529,951
CC FILING DATE: May 29, 1990
CC APPLICATION NUMBER: US 07/325,365
CC FILING DATE: March 17, 1989
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Amy E. Mandragoras
CC REGISTRATION NUMBER: 36,207
CC REFERENCE/DOCKET NUMBER: IM-018DV
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617)227-7400
CC TELEFAX: (617)227-5941
CC INFORMATION FOR SEQ ID NO: 75:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 397 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 76 AA: 42947 MW: 781425 CN:
Query Match 13.7% Score 141 DB 1 Length 397
Best Local Similarity 37.9% Pred. No. 2,06e-03
Matches 22: Conservative 12: Mismatches 22: Indels 2: Gaps 2:
DE 108 FAAGNRPLWLIIFKNDYVNL ELVNSDKIIDRGYKVE-I-INGGLILMNVKNI 163
QY 67 FTLMGRPLWLIIFSGNNIKLMPMTIAGYKTFDGRRAEVSVHVHGAKFIRRVGSI 124
RESULT 5
ID US-08-290-448A-76 STANDARD: PRT: 397 AA.
XX
AC xxxxxx
DE
XX
DE Sequence 76, Application US/08290448A
CC
CC Sequence 76, Application US/08290448A
CC Patent No. 5676954
CC GENERAL INFORMATION:
CC APPLICANT: Rogers, Bruce
CC APPLICANT: Klapper, David G.
CC APPLICANT: Rafnar, Thorunn
CC APPLICANT: Kuo, Mei-chang
CC TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses
CC NUMBER OF SEQUENCES: 93

CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: LAHIVE & COCKFIELD
CC STREET: 60 State Street, suite 510
CC CITY: Boston
CC STATE: Massachusetts
CC COUNTRY: USA
CC ZIP: 02109-1875
CC
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent in Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/290-448A
CC FILING DATE: August 15, 1994
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/529,951
CC FILING DATE: May 29, 1990
CC APPLICATION NUMBER: US 07/325,365
CC FILING DATE: March 17, 1989
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Amy E. Mandragoras
CC REGISTRATION NUMBER: 36,207
CC REFERENCE/DOCKET NUMBER: IM-018CN
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617)227-7400
CC TELEFAX: (617)227-5941
CC INFORMATION FOR SEQ ID NO: 76:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 397 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 397 AA: 42947 MW: 781425 CN:
Query Match 13.7% Score 141 DB 1 Length 397
Best Local Similarity 37.9% Pred. No. 2,06e-03
Matches 22: Conservative 12: Mismatches 22: Indels 2: Gaps 2:
DB 108 FAAGNRPLWLIIFKNDYVNL ELVNSDKIIDRGYKVE-I-INGGLILMNVKNI 163
QY 67 FTLMGRPLWLIIFSGNNIKLMPMTIAGYKTFDGRRAEVSVHVHGAKFIRRVGSI 124
RESULT 6
ID US-08-290-448A-76 STANDARD: PRT: 397 AA.
XX
AC xxxxxx
DE
XX
DE Sequence 76, Application US/08290448A
CC
CC Sequence 76, Application US/08290448A
CC Patent No. 5698204
CC GENERAL INFORMATION:
CC APPLICANT: Rogers, Bruce
CC APPLICANT: Klapper, David G.
CC APPLICANT: Rafnar, Thorunn
CC APPLICANT: Kuo, Mei-chang
CC TITLE OF INVENTION: Allergenic Proteins From Ragweed and Uses
CC NUMBER OF SEQUENCES: 93
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: LAHIVE & COCKFIELD
CC STREET: 60 State Street, suite 510
CC CITY: Boston
CC STATE: Massachusetts
CC COUNTRY: USA
CC ZIP: 02109-1875
CC
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: Patent Release #1.0, Version #1.25
 CC CURRENT APPLICATION DATA:
 CC APPLICATION NUMBER: US/08/290,448A
 CC FILING DATE: August 15, 1994
 CC PRIOR APPLICATION DATA:
 CC APPLICATION NUMBER: US 07/529,951
 CC FILING DATE: May 29, 1990
 CC APPLICATION NUMBER: US 07/325,365
 CC FILING DATE: March 17, 1989
 CC ATTORNEY/AGENT INFORMATION:
 CC NAME: Amy E. Mandragoras
 CC REGISTRATION NUMBER: 36,207
 CC REFERENCE/DOCKET NUMBER: IM1-C18CN
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (617)227-5941
 CC TELEFAX: (617)227-7400
 CC INFORMATION FOR SEQ ID NO: 163:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 391 amino acids
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC SEQUENCE: 391 AA: 42864 MW: 761425.0N

Query Match: 13.4% Score 147. DB 1. Length 197
 Best Local Similarity: 39.7% Ref No. 4316 33 Index 20 Gaps 20
 Matches 23: Conservative 12: Mismatches 12: Indels 2: Gaps 2:
 DB 102 FAAGNRRLPIWIFKSNMVIHLNDELVNSKTIIDGRGVKVAEILINGULTMNVKNI:163
 QY 67 FILMGRPLKLIFFSNMNIKMKMPMEIAGYKTFDGRPAEVSIVVNGAKFPRVDGI:124

RESULT 8
 ID US-08-175-809A-59 STANDARD: PRI: 391 AA
 AC xxxxxx
 XX
 XX

Sequence 59, Application US/08/290,448A

Sequence 59, Application US/08/290,448A
 Patent No. 5676354
 GENERAL INFORMATION:
 APPLICANT: Rogers, Bruce
 APPLICANT: Klapper, David G.
 APPLICANT: Rafnar, Thorunn
 APPLICANT: Kuo, Mel-chand
 TITLE OF INVENTION: Atherosclerotic Protein From Bacteria and Uses
 NUMBER OF SEQUENCES: 33
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: LAHIVE & CUCKFIELD, LLP
 STREET: 60 State Street
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02109-1875
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/290,448A
 FILING DATE: August 15, 1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/529,951
 FILING DATE: May 29, 1990
 APPLICATION NUMBER: US 07/325,365
 FILING DATE: March 17, 1989
 ATTORNEY/AGENT INFORMATION:
 NAME: Amy E. Mandragoras
 REGISTRATION NUMBER: 36,207
 REFERENCE/DOCKET NUMBER: IM1-C18CN
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617)227-5941
 TELEFAX: (617)227-7400
 INFORMATION FOR SEQ ID NO: 163:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 391 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein

CC REGISTRATION NUMBER: 36,207
 CC REFERENCE/DOCKET NUMBER: IM1-C18CN
 CC TELECOMMUNICATION INFORMATION:
 CC TELEPHONE: (617)227-5941
 CC TELEFAX: (617)227-7400
 CC INFORMATION FOR SEQ ID NO: 59:
 CC SEQUENCE CHARACTERISTICS:
 CC LENGTH: 391 amino acid
 CC TYPE: amino acid
 CC TOPOLOGY: linear
 CC MOLECULE TYPE: protein
 CC SEQUENCE: 391 AA: 42864 MW: 761425.0N

Query Match: 13.4% Score 147. DB 1. Length 197
 Best Local Similarity: 39.7% Ref No. 4316 33 Index 20 Gaps 20
 Matches 23: Conservative 12: Mismatches 12: Indels 2: Gaps 2:
 DB 102 FAAGNRRLPIWIFKSNMVIHLNDELVNSKTIIDGRGVKVAEILINGULTMNVKNI:163
 QY 67 FILMGRPLKLIFFSNMNIKMKMPMEIAGYKTFDGRPAEVSIVVNGAKFPRVDGI:124

RESULT 8
 ID US-08-290-448A-59 STANDARD: PRI: 391 AA
 AC xxxxxx
 XX
 XX

Sequence 59, Application US/08/290,448A

Sequence 59, Application US/08/290,448A
 Patent No. 5676354
 GENERAL INFORMATION:
 APPLICANT: Rogers, Bruce
 APPLICANT: Klapper, David G.
 APPLICANT: Rafnar, Thorunn
 APPLICANT: Kuo, Mel-chand
 TITLE OF INVENTION: Atherosclerotic Protein From Bacteria and Uses
 NUMBER OF SEQUENCES: 33
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: LAHIVE & CUCKFIELD, LLP
 STREET: 60 State Street, suite 510
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02109-1875
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/290,448A
 FILING DATE: August 15, 1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/529,951
 FILING DATE: May 29, 1990
 APPLICATION NUMBER: US 07/325,365
 FILING DATE: March 17, 1989
 ATTORNEY/AGENT INFORMATION:
 NAME: Amy E. Mandragoras
 REGISTRATION NUMBER: 36,207
 REFERENCE/DOCKET NUMBER: IM1-C18CN
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617)227-5941
 TELEFAX: (617)227-7400
 INFORMATION FOR SEQ ID NO: 59:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 391 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein

```

RESULT 10
ID US-8-175-069A-74 STANDARD 10 PRT: 398 AA.
XX
XX xxxxxx
XX
XX
XX
XX
XX
XX
Sequence 74, Application US/8175-069A
CC
CC Sequence 74, Application US/8175-069A
CC Patent No. 576761
CC GENERAL INFORMATION:
CC APPLICANT: Rogers, Brian
CC APPLICANT: Klapper, David G.
CC APPLICANT: Rainier, Thomas J.
CC APPLICANT: Koo, Mel-Chen
CC TITLE OF INVENTION: Amino Acid Proteins from Rainbow Trout
CC NUMBER OF SEQUENCES: 9
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: LAHIVE & FELD, LLP
CC STREET: 60 State Street
CC CITY: Boston
CC STATE: Massachusetts
CC COUNTRY: USA
CC ZIP: 02109-1875
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: Patent Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/8175-069A
CC FILING DATE: December 29, 1993
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: US 07/529,951
CC FILING DATE: May 29, 1995
CC APPLICATION NUMBER: US 07/325,365
CC FILING DATE: March 17, 1989
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Amy E. Mandragoras
CC REGISTRATION NUMBER: 46,207
CC REFERENCE/DOCKET NUMBER: 2M1-018CV
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617)227-7400
CC TELEFAX: (617)227-5974
CC INFORMATION FOR SEQ ID NO. 74:
CC SEQUENCE CHARACTERISTICS
CC LENGTH: 398 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: protein
CC SEQUENCE 398 AA: 43664 MW: 774006 CN:
Query Match: 13.4% Score 137, DB 1, Length 398;
Best Local Similarity 39.7%, Pred. No. 4, 3le-03;
Matches 23; Conservative 10; Mismatches 23; Indels 2; Gaps
DB 109 FAAGNRPRLHIKRNHV:HLNDLWVNSDKTIDGGGVKNVIINA-GLTEN-VANII 164
QY : ||||| | | | | | | | | | | | | | | | | | | | | |
67 FTLMGRPLWIIFSGNN:NKLKMPMYIAGVKTFDGRRAEVSYYHHNGAKFIKKVDGI: 124

RESULT 11
ID JS-08-290-448A-74 STANDARD: PRT: 398 AA.
XX
XX xxxxxx
XX
XX
XX
XX
XX
XX
Sequence 74, Application US/.8290448A
CC
CC Sequence 74, Application US/.8290448A
CC Patent No. 5676954
CC

```


Matches 20 Conservative 10 Mismatches 24 Indels 1 Gaps 1
59 59AAGAGGRLWIFERLWVRLDRELAINNDKTDGRGAKVELINAGFAIYNVKNI 153
|||||
67 67LILMGRRRLWIFESNNIKKKRPMYIAGYKTFGRRAEVSYYEVNGAKF-IRPV 120

Search completed: Tue Jun 20 14:33:43 2000
Run time: 1.4 secs.

CC ATTORNEY/AGENT INFORMATION:
CC NAME: Darlene A. Vanstone
CC REGISTRATION NUMBER: 35,729
CC REFERENCE/DOCKET NUMBER: 025.5 US (IMI-028CP2)
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 466-6000
CC TELEFAX: (617) 466-6040
CC INFORMATION FOR SEQ ID NO: 62:
CC LENGTH: 60 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC FRAGMENT TYPE: internal
CC SEQUENCE 60 AA: 6644 MW: 19454.0N

CC Query Match 28.8% Score 296: 59.8: Length 60:
CC Best Local Similarity 73.5% Pred. No. 3,14e-18:
CC Matches 39: Conservative 5: Mismatches 7: Indels 2: Gaps 2:

DE 6 RPLWIFSGNMN:KLMKPMYIAGYKTFDGRGAQV-YVGNQGPVWEIKRVSNV: 57
QY 73 RPLWIFSGNMN:KLMKPMYIAGYKTFDGRPAFVSVYVHNGA-KFIRRYDGI: 124

RESULT 5
ID US-09-142-524A-62 STANDARD: PRT: 60 AA.
AC xxxxxx
XX
DT
XX
DE Sequence 62, Application US/08226248A
XX
CC Sequence 62, Application US/08226248A
CC GENERAL INFORMATION:
CC APPLICANT: Griffith, Irwin J.
CC APPLICANT: Pofflock, Joanne
CC APPLICANT: Bord, Julian F.
CC APPLICANT: Garman, Richard D.
CC APPLICANT: Kuo, Mei-Chang
CC APPLICANT: Yeung, Stephen H.
CC APPLICANT: Exley, Mark A.
CC APPLICANT: Powers, Steven P.
CC APPLICANT: Brauer, Andrew
CC APPLICANT: Exley, Mark A.
CC APPLICANT: Powers, Steven P.
CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
CC TITLE OF INVENTION: Japanese Cedar Pollen
CC NUMBER OF SEQUENCES: 20
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
CC STREET: 610 Lincoln St
CC CITY: Waltham
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02154
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/226,248A
CC FILING DATE: April 8, 1994
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 07/938,990
CC FILING DATE: September 1, 1992
CC APPLICATION NUMBER: PCT/US93/00139
CC FILING DATE: January 15, 1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Darlene A. Vanstone
CC REGISTRATION NUMBER: 35,729
CC REFERENCE/DOCKET NUMBER: 025.5 US (IMI-028CP)

CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 466-6000
CC TELEFAX: (617) 466-6040
CC INFORMATION FOR SEQ ID NO: 62:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 60 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC MOLECULE TYPE: peptide
CC FRAGMENT TYPE: internal
CC SEQUENCE 60 AA: 6644 MW: 19454.0N

CC Query Match 28.8% Score 296: 59.8: Length 60:
CC Best Local Similarity 73.5% Pred. No. 3,14e-18:
CC Matches 39: Conservative 5: Mismatches 7: Indels 2: Gaps 2:

DE 6 RPLWIFSGNMN:KLMKPMYIAGYKTFDGRGAQV-YVGNQGPVWEIKRVSNV: 57
QY 73 RPLWIFSGNMN:KLMKPMYIAGYKTFDGRPAFVSVYVHNGA-KFIRRYDGI: 124

RESULT 6
ID US-08-467-006-62 STANDARD: PRT: 60 AA.
AC xxxxxx
XX
DT
XX
DE Sequence 62, Application US/ 8457006
XX
CC Sequence 62, Application US/ 8457006
CC GENERAL INFORMATION:
CC APPLICANT: Griffith, Irwin J.
CC APPLICANT: Pofflock, Joanne
CC APPLICANT: Bord, Julian F.
CC APPLICANT: Garman, Richard D.
CC APPLICANT: Kuo, Mei-Chang
CC APPLICANT: Yeung, Stephen H.
CC APPLICANT: Exley, Andrew
CC APPLICANT: Exley, Mark A.
CC APPLICANT: Powers, Steven P.
CC TITLE OF INVENTION: Allergenic Proteins And Peptides From
CC TITLE OF INVENTION: Japanese Cedar Pollen
CC NUMBER OF SEQUENCES: 20
CC CORRESPONDENCE ADDRESS:
CC ADDRESSEE: Immunologic Pharmaceutical Corporation, Inc.
CC STREET: 610 Lincoln St
CC CITY: Waltham
CC STATE: MA
CC COUNTRY: USA
CC ZIP: 02154
CC COMPUTER READABLE FORM:
CC MEDIUM TYPE: Floppy disk
CC COMPUTER: IBM PC compatible
CC OPERATING SYSTEM: PC-DOS/MS-DOS
CC SOFTWARE: PatentIn Release #1.0, Version #1.25
CC CURRENT APPLICATION DATA:
CC APPLICATION NUMBER: US/08/467,006
CC FILING DATE: June 5, 1995
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/350,225
CC FILING DATE: December 6, 1994
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Jane E. Remillard
CC REGISTRATION NUMBER: 38,872
CC REFERENCE/DOCKET NUMBER: 025.6 USD6 (IMI-028CPD6)
CC TELECOMMUNICATION INFORMATION:
CC TELEPHONE: (617) 227-7400
CC TELEFAX: (617) 227-5941
CC INFORMATION FOR SEQ ID NO: 62:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 60 amino acids


```

RE-ULT 9
ID US-08-468-940-62 STANDARD: PRT: 60 AA.
XX
AC
XX
XX
DT
XX
DE
XX
Sequence 62, Application: US/08468940
Sequence 62, Application: US/08468940
GENERAL INFORMATION:
APPLICANT: Griffiths, Irwin J.
APPLICANT: Pollock, Joanne
APPLICANT: Bond, Julian F.
APPLICANT: Gorman, Richard D.
APPLICANT: Kuo, Mei-Chang
APPLICANT: Yeung, Siu-mei H.
APPLICANT: Brauer, Andrew
APPLICANT: Exley, Mark A.
APPLICANT: Powers, Steven P.
TITLE OF INVENTION: Allergenic Proteins And Peptides From
TITLE OF INVENTION: Japanese Cedar Pollen
NUMBER OF SEQUENCES: 261
CORRESPONDENCE ADDRESS:
ADDRESS: Immunologic Pharmaceutical Corporation, Inc.
STREET: 610 Lincoln St
CITY: Waltham
STATE: MA
COUNTRY: USA
ZIP: 02154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,940
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/450,225
FILING DATE: December 6, 1994
APPLICATION NUMBER: 08/226,248
FILING DATE: April 8, 1994
APPLICATION NUMBER: 07/938,990
FILING DATE: September 1, 1992
APPLICATION NUMBER: PC/US93/00139
FILING DATE: January 15, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Parlese A. Vassalopoulos
REGISTRATION NUMBER: 35,724
REFERENCE/DOCKET NUMBER: 025.6 US (IM-228CP2)
INTELLIGENCE INFORMATION:
TELEPHONE: (617) 466-6000
TELEFAX: (617) 466-6040
INFORMATION FOR SEQ ID NO: 62:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
SEQUENCE 50 AA: 6644 MW: 19464 CN:
Query Match: 28.8% Score 296: DB 10: Length 60:
Best Local Similarity 73.6% Pred. No. 3,14e-18;
Matches 39: Conservative 5: Mismatches 7: Indels 2: Gaps 2:
DE 6 RPLWIFSGNMNKLKMPYIAGYKTFDGRGACV-YINGGPGCVFKRVSNVI 57
QY 73 RPLWIFSGNMNKLKMPYIAGYKTFDGRGACV-YINGGPGCVFKRVSNVI 124

```

```

RE-ULT 10
ID US-08-467-023-62 STANDARD: PRT: 60 AA.
XX
AC
XX
XX
DT
XX
DE
XX
Sequence 62, Application: US/08467023
Sequence 62, Application: US/08467023
GENERAL INFORMATION:
APPLICANT: Griffiths, Irwin J.
APPLICANT: Pollock, Joanne
APPLICANT: Bond, Julian F.
APPLICANT: Gorman, Richard D.
APPLICANT: Kuo, Mei-Chang
APPLICANT: Yeung, Siu-mei H.
APPLICANT: Brauer, Andrew
APPLICANT: Exley, Mark A.
APPLICANT: Powers, Steven P.
TITLE OF INVENTION: Allergenic Proteins And Peptides From
TITLE OF INVENTION: Japanese Cedar Pollen
NUMBER OF SEQUENCES: 261
CORRESPONDENCE ADDRESS:
ADDRESS: Immunologic Pharmaceutical Corporation, Inc.
STREET: 610 Lincoln St
CITY: Waltham
STATE: MA
COUNTRY: USA
ZIP: 02154
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/467,023
FILING DATE: June 6, 1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/450,225
FILING DATE: December 6, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Jane E. Remillard
REGISTRATION NUMBER: 38,872
REFERENCE/DOCKET NUMBER: 025.6 USC2 (IM-028CPD2)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7426
TELEFAX: (617) 227-6941
INFORMATION FOR SEQ ID NO: 62:
SEQUENCE CHARACTERISTICS:
LENGTH: 60 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
SEQUENCE 50 AA: 6644 MW: 19464 CN:
Query Match: 28.8% Score 296: DB 10: Length 60:
Best Local Similarity 73.6% Pred. No. 3,14e-18;
Matches 39: Conservative 5: Mismatches 7: Indels 2: Gaps 2:
DB 6 RPLWIFSGNMNKLKMPYIAGYKTFDGRGACV-YINGGPGCVFKRVSNVI 57
QY 73 RPLWIFSGNMNKLKMPYIAGYKTFDGRGACV-YINGGPGCVFKRVSNVI 124

```


CC APPLICATION NUMBER: US/08/350.225
CC FILING DATE: December 6, 1994
CC CLASSIFICATION: 424
CC PRIOR APPLICATION DATA:
CC APPLICATION NUMBER: 08/226.248
CC FILING DATE: April 8, 1994
CC APPLICATION NUMBER: 07/538.990
CC FILING DATE: September 1, 1992
CC APPLICATION NUMBER: PCI/US93/00139
CC FILING DATE: January 15, 1993
CC ATTORNEY/AGENT INFORMATION:
CC NAME: Carlene A. Vanstone
CC REGISTRATION NUMBER: 35.725
CC REFERENCE/DOCKET NUMBER: 025.6 US (IM-C28CP2)
CC INFORMATION INFORMATION:
CC TELEPHONE: (617) 466-6000
CC TELEFAX: (617) 466-6040
CC INFORMATION FOR SEQ ID NO: 2:
CC SEQUENCE CHARACTERISTICS:
CC LENGTH: 374 amino acids
CC TYPE: amino acid
CC TOPOLOGY: linear
CC Molecule type: protein
CC SEQUENCE 374 AA: 40220 MW: 75747.7 CN:

Query Match 28.8% Score 296 DB 3 Length 374
Best Local Similarity 73.6% Freq. NO. 3.14e-18:
Matches 39: Conservative 5: Mismatches 7: Indels 2: Gaps 2:

18 87 RQKIIIFSCNMNIKMKMPNYIAGYKTFDQSGAV-VIGNSGPCVFIRKRVSNV:138
19 73 RQKIIIFSCNMNIKMKMPNYIAGYKTFDQSGAV-VIGNSGPCVFIRKRVSNV:124

Search completed: Mon Jun 19 16:21:59 2000
Run time: 44 secs.


```

26-Apr-1999
J02123: PC2065
REFERENCE
J02123
Authors
Koi Kito, K.
Journal
Biochem Biophys Res Commun. (1994) 199:619-625
Title
Cloning and sequencing of cDNA coding for Crp J I, a major
allergen of Japanese cedar pollen.
Cross-References M01034:59234
Accession
J02123
Molecule_type mRNA
Residues 1-174 #label SGN
Cross-References GB:026544: NID-0433631: PID:01006085: PID:04943632
Experimental_source pollen
Accession PC2065
Molecule_type protein
Residues 22-53:58-91:219-232:236-248:259-307:345-372 #label SC2
#note
the authors described car: hydrate binding site for
residue 275
CLASSIFICATION
#superfamily pectate lyase LA159
KEYWORDS
#glycoprotein: pollen
FEATURE
1-21
#domain signal sequence #status predicted #label SGN
#product major allergen Crp J I (clone PC2-2-2) #status
predicted #label MA1N
198-199:233-354 #binding site carbohydrate (ASN) (covariant) #status
predicted
SUMMARY
#length 374 #molecular-weight 40645 #checksum 2926
Query Match 28.8% Score 296 DB 2: Length 374:
Best Local Similarity 73.6% Pred. No: 4,956-53:
Matches 39: Conservative 5: Mismatches 7: Indels 2: Gaps 2:
Dt 87 RPWKIIFFSNMNIKKMPYIAGYKTFDGRGACV-VIOMGGPCVFIKQVSNV 138
RPWKIIFFSNMNIKKMPYIAGYKTFDGRRAEVSIVHNGA-KFIRVCGII 124
QY 73 RPWKIIFFSNMNIKKMPYIAGYKTFDGRRAEVSIVHNGA-KFIRVCGII 124
RPWKIIFFSNMNIKKMPYIAGYKTFDGRRAEVSIVHNGA-KFIRVCGII 124
RESULT 3
ENTRY
J02123 #type complete
#molecule_type mRNA
#residues 1-174 #label KOM
#cross-References M01034:59234: NID:0506657: PID:01006734: PID:0506658
#accession PC2145
#molecule_type protein
#residues 52-61 #label KO2
REFERENCE
A60147
Authors
Sakaguchi, M., Ino, Y., Saito, T., Tanaka, M., Ando, S., US, I., Miki,
M., Matsuda, T.
Allergy (1993) 45:309-312
Title
Identification of the second major allergen of Japanese cedar
pollen.
Cross-References M01034:59238
Accession A60147
Molecule_type protein
Residues 55-64 #label SAK
#cross-References GB:026544: NID-0433631: PID:01006085: PID:04943632
Experimental_source pollen
KEYWORDS
#glycoprotein: pollen
FEATURE
1-54
#domain signal sequence #status predicted #label SGN
#product second major allergen Crp J #status predicted
#label MA1N

```

```

429,450,472 #binding site carbohydrate (ASN) (covariant) #status
predicted
SUMMARY
#length 514 #molecular-weight 5648 #checksum 472
Query Match 16.6% Score 172 DB 2: Length 514:
Best Local Similarity 95.2% Pred. No: 1,248-11:
Matches 20: Conservative 1: Mismatches 0: Indels 0: Gaps 0:
Db 133 WKNNRIWQFAKLTGFLMKR 159
III III III III III III
QY 52 WKNNRIWQFAKLTGFLMKR 72
WKNNRIWQFAKLTGFLMKR 72
RESULT 4
ENTRY
S48730 #type complete
#molecule_type mRNA
#residues 1-514 #label NAM
#cross-References GR:037755: NID:0577655: PID:01007022: PID:0577656
#accession S48730
#status preliminary
#journal
Fukuda, S., Usui, M., Kaitoh, M.
PES Lett. (1994) 45:124-128
Title
Molecular cloning of the second major allergen Crp J I
from Japanese cedar pollen
Cross-References M01034:591077
Accession S48730
#domain signal sequence #status predicted #label SGN
#product major allergen Crp J I (clone PC2-2-2) #status
predicted #label MA1N
198-199:233-354 #binding site carbohydrate (ASN) (covariant) #status
predicted
SUMMARY
#length 514 #molecular-weight 5645 #checksum 472
Query Match 16.8% Score 172 DB 2: Length 514:
Best Local Similarity 95.2% Pred. No: 1,248-11:
Matches 20: Conservative 1: Mismatches 0: Indels 0: Gaps 0:
Db 133 WKNNRIWQFAKLTGFLMKR 159
III III III III III III
QY 52 WKNNRIWQFAKLTGFLMKR 72
WKNNRIWQFAKLTGFLMKR 72
RESULT 5
ENTRY
T06728 #type complete
#molecule_type mRNA
#residues 1-514 #label CDF
#cross-References M01034:591077: NID:0577655: PID:01007022: PID:0577656
#accession T06728
#status preliminary
#journal
Quatier, F., Choisme, N., Robert, C., Brothier, P., Wilkner,
P., Catala, J., Artigues, P., Salas, M.,
Weissenbach, J., Salas, M.,
K.F., J. Scheller, C.
Submitted to the Protein Sequence Database, April 1999
Cross-References M01034:591077: NID:0577655: PID:01007022: PID:0577656
Accession T06728
#domain signal sequence #status predicted #label SGN
#product second major allergen Crp J #status predicted
#label MA1N
SUMMARY
#length 514 #molecular-weight 5645 #checksum 472
Query Match 15.1% Score 155 DB 2: Length 514:

```



```

Best Local Similarity 57.6%; Pred. No. 6,01e-09;
Matches 19; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

DB 146 PLWVFSNNMLIRKQELIINSYKTIQGRSAV 178
      |||||  |||||  |||||  |||||  |||||
QY 74 PMITFSNNMIKIKMPMYAGYKTFDGRRAEV 106

RESULT 6
ENTRY pectate lyase (EC 4.2.2.2) - common tobacco
TITLE pectate lyase (EC 4.2.2.2) - common tobacco
CKANISM #forma_name Nicotiana glauca common tobacco
DATE 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change
29-Sep-1999
ACCESSIONS S26211 S26212 S21933 S22753 S22954
REFERENCE S26211
#authors Rogers, H.J.; Harvey, A.; Lonsdale, D.M.
#journal Plant Mol. Biol. (1992) 20:493-502
#title Isolation and characterization of a tobacco gene with
#title homology to pectate lyase which is specifically expressed
#title during microsporogenesis.
#cross-references MIM:31043035
#accession S26211
#molecule_type DNA
#residues 1-397 #label MOG
#cross-references EMBL:X67158; NID:g19937; PIDN:CAA47630.1; PID:g19908
#accession S26212
#molecule_type mRNA
#residues 119-155; C 157-188; G 191-199; D 200; R 203-248;
#cross-references EMBL:X67159
#note Translation of the nucleotide sequence is not complete
REFERENCE S21933
#authors Lonsdale, D.M.
#submission Submitted to the EMBL Data Library, July 1991
#accession S21933
#molecule_type DNA
#residues 1-397 #label LON
#cross-references EMBL:X61102; NID:g19981; PIDN:CAA43414.1; PID:g19982
GENETICS 193/1: 293/2
#initiator superfamily pectate lyase LAT59
CLASSIFICATION carbon-oxygen lyase
KEYWORDS #length 397 #molecular-weight 44351 #checksum 1480
SUMMARY
Query Match 11.8%; Score 152; DB 2; Length 397;
Best Local Similarity 36.7%; Pred. No. 1.75e-08;
Matches 22; Conservative 15; Mismatches 22; Indels 1; Gaps 1;

DB 105 YGVICRPLWIFGNMKIKUSRELIVSNKTIQGRGFNVHUNAGTKTQASNTIIS 163
      |||||  |||||  |||||  |||||  |||||
QY 67 FTLMGRPLWIFSGNNIKIKMPMYIAYKTFDGRRAEVSVHVNAGKTRVVGIIAA 126

RESULT 7
ENTRY pectate lyase (EC 4.2.2.2) - trumpet lily
TITLE pectate lyase (EC 4.2.2.2) - trumpet lily
CKANISM #forma_name Lilium longiflorum #common_name trumpet lily
DATE 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change
29-Sep-1999
ACCESSIONS S29612
REFERENCE S29611
#authors Kim, S.; Finkel, D.J.; An, G.
#submission Submitted to the EMBL Data Library, October 1992
#description Abundance patterns of lily pollen cDNAs: characterization of
#description three pollen-preferential cDNA clones.
#accession S29612
#status preliminary
#molecule_type mRNA
#residues 1-434 #label KIM
#cross-references EMBL:Z17338; NID:g19450; PIDN:CAA78976.1; PID:g19451
#experimental_source cv. Nellie White, mature flower
CLASSIFICATION #superfamily pectate lyase LAT59

KEYWORDS
SUMMARY
Query Match 13.7%; Score 141; DB 2; Length 396;
Best Local Similarity 43.2%; Pred. No. 8.25e-07;
Matches 19; Conservative 8; Mismatches 17; Indels 0; Gaps 0;

DB 107 FGAAQRPLWIFERDMVIRLQKENVVNSDKTIDGKAGKEIIN 150
      |||||  |||||  |||||  |||||  |||||
QY 67 FTLMGRPLWIFSGNNIKIKMPMYIAYKTFDGRRAEVSVHVN 110

KEYWORDS
SUMMARY
Query Match 13.9%; Score 143; DB 2; Length 404;
Best Local Similarity 39.5%; Pred. No. 4.13e-07;
Matches 25; Conservative 14; Mismatches 22; Indels 4; Gaps 4;

DB 140 FQTLR-WGVICRPLWIFGKSMVRLKCELLINNDKTIQGRJANVO-TFADMAQIVVF 199
      |||||  |||||  |||||  |||||  |||||
QY 61 FAKLTGFTLMGRPLWIFSGNNIKIKMPMYIAYKTFDGRRAEVSVHVNQAK 194
      |||||  |||||  |||||  |||||  |||||
DB 197 VHNV: 201
QY 120 VDGII 124

RESULT 8
ENTRY #type complete
TITLE allergen Amb a 1 precursor - common ragweed
CKANISM #forma_name Artemisia artemisiifolia #common_name ragweed
DATE 27-Nov-1991 #sequence_revision 03-Apr-1992 #text_change
29-Sep-1999
ACCESSIONS A39099; A60895; A53240
REFERENCE A39099
#authors Rainey, T.; Griffith, L.J.; Kuo, M.; Bond, J.F.; Roberts,
#journal J. Biol. Chem. (1991) 266:1229-1236
#title Cloning of Amb a 1 (allergen E), the major allergen family of
#title short ragweed pollen.
#cross-references MIM:3103235
#accession A39099
#molecule_type mRNA
#residues 1-396 #label RAF
#cross-references GIM:53116
#accession A60895
#authors Smith, J.J.; Olson, J.M.; Klapper, E.G.
#journal Mol. Immunol. (1988) 25:155-165
#title Monoclonal antibodies to denatured ragweed pollen allergen
#title Amb a 1: characterization, specificity for the denatured
#title allergen, and utilization for the isolation of immunogenic
#title peptides of Amb a 1.
#cross-references MIM:3103235
#accession A60895
#molecule_type protein
#residues 256-273; 252-303; W 305-306 #label SM1
REFERENCE A53240
#authors Griffith, L.J.; Clark, J.; Klapper, E.G.; Roberts, P.L.;
#journal Int. Arch. Allergy Appl. Immunol. (1992) 9: 299-314
#title Sequence polymorphism of Amb a 1 and Amb a 2 in the major
#title allergens in Artemisia artemisiifolia (Sant. Rudbeck)
#cross-references MIM:3103235
#accession A53240
#status preliminary
#molecule_type mRNA
#residues 1-396 #label GR1
#cross-references GIM:53058; NID:g166434; PIDN:AAA3235.1; PID:g166435
CLASSIFICATION #superfamily pectate lyase LAT59
KEYWORDS glycoprotein; pollen
FEATURE
36 #binding-site carbohydrate (Asn) (covalent) #status
predicted
SUMMARY
Query Match 13.7%; Score 141; DB 2; Length 396;
Best Local Similarity 43.2%; Pred. No. 8.25e-07;
Matches 19; Conservative 8; Mismatches 17; Indels 0; Gaps 0;

DB 107 FGAAQRPLWIFERDMVIRLQKENVVNSDKTIDGKAGKEIIN 150
      |||||  |||||  |||||  |||||  |||||
QY 67 FTLMGRPLWIFSGNNIKIKMPMYIAYKTFDGRRAEVSVHVN 110

```


29-Sep-1999
 ACCESSIONS 107058: SC8383
 REFERENCE 215890
 #authors King, R.A.
 #submission Submitted to the EMBL Data Library, September 1994
 #accession 107058
 #status translated from GR/EMBL/DDP
 #molecule_type DNA
 #label WIN
 #residues 1-398
 #cross-references EMBL:X15500; NID:p551555; P1DN:CAA33524.1;
 P1D:p551556
 #experimental_source strain VF36; another

REFERENCE SC8383
 #authors King, R.A.; Yaraaguchi, J.; Larabell, S.K.; Ursin, V.M.;
 McCormick, S.
 #journal Plant Mol. Biol. (1989) 14:17-28
 #title Molecular and genetic characterization of two
 pollen-expressed genes that have sequence similarity to
 pectate lyases of the plant pathogen Erwinia.
 #cross-references M01D:91322485
 #accession SC8383
 #molecule_type DNA
 #residues 1-241; 26-49; 51-61; 63-66; 67-69-215; 217-
 118-141; 143-166; 168-200; 202-215; 217-349; 351-371;
 373-383; 385-398; #label WIN
 #cross-references EMBL:X15500

GENETICS

#map_position 3
 #introns 194/1; 294/2
 #superfamily pectate lyase LAT55
 #carbon-oxygen lyase
 #length 398 #molecular-weight 44563 #checksum 5266
 Query Match 13.4% Score 136; DB 2; Length 398;
 Best Local Similarity 35.5% Pred. No. 2; 1e-06;
 Matches 22; Conservative 17; Mismatches 5; Gaps 4;

DB 106 EVUQKPIWTFARNRRLTRELIVSNKTIIDGR-GK--YVHANGAGIKQASNYI 162
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 QY 57 FTLMGRRLWII:FSGNMNIKLKKMYPYIAGYKTFGGRRAEV 105
 DB 163 IS 164
 QY 125 AA 126

RESULT 14

ENTRY B39099 #type complete
 TITLE Allergen Amb a 1 - common ragweed
 ORGANISM Ambrosia artemisiifolia #submitter DATE 04-09-99
 #reviewed
 DATE 27-Nov-1999 #sequence_revision 01-Apr-1992 #text_change 18-Sep 1999

ACCESSIONS B39099
 REFERENCE B39099
 #authors Kainer, T.; Griffith, T.C.; Kuo, M.; Bond, J.F.; Rogers,
 B.L.; Klapper, D.G.
 #journal J. Biol. Chem. (1991) 266:1229-1236
 #title Cloning of Amb a 1 (anion. 5), the major allergen family of
 short ragweed pollen
 #cross-references M01D:9105325

#accession B39099
 #status Preliminary
 #molecule_type mRNA
 #residues 1-338 #label RAF
 #superfamily pectate lyase LAT59
 #carbon-oxygen lyase
 #length 398 #molecular-weight 43664 #checksum 417

Query Match 13.4% Score 137; DB 2; Length 398;
 Best Local Similarity 39.7% Pred. No. 3; 26e-06;
 Matches 23; Conservative 10; Mismatches 23; Indels 2; Gaps 2;

DB 109 FAAQNRPLIIFKRNVMVIRHNLQELVNSDKTIIDGRGVKVNIVNA-GLTLKN-VKNII 164
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 QY 57 FTLMGRRLWII:FSGNMNIKLKKMYPYIAGYKTFGGRRAEVSVHVNDKAFIRVAGII 124
 RESULT 15
 ENTRY #type complete
 TITLE pectate lyase (EC 4.2.2.2) - tomato
 ORGANISM Lycopersicon esculentum #submitter DATE 19-Mar-1997
 #sequence_revision 19-Mar-1997 #text_change 23-Jul-1999

ACCESSIONS S12209
 REFERENCE S12209
 #authors Budellier, K.A.; Smith, A.G.; Gasser, C.S.
 #journal Mol. Gen. Genet. (1993) 224:183-192
 #title Regulation of a nuclear transmitting tissue-specific gene in
 wild-type and tumorigenic tomato and tobacco.
 #cross-references M01D:9111718
 #accession S12209
 #status Preliminary

#molecule_type mRNA
 #residues 1-404 #label QD
 #cross-references GB:X55133; NID:p19161; P1DN:CAA48479.1; F101919.62
 #superfamily pectate lyase LAT59
 #carbon-oxygen lyase
 #length 404 #molecular-weight 44296 #checksum 4917

Query Match 13.4% Score 137; DB 2; Length 404;
 Best Local Similarity 54.5% Pred. No. 3; 26e-06;
 Matches 18; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

DB 120 PLWIFKRMVQLKQELVMSYKTIIDGRGASV 152
 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 QY 74 PLWIFSGNMNIKLKKMYPYIAGYKTFGGRRAEV 105

Search completed: Mon Jun 19 15:58:21 2000
 Job time: 17 secs.

 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100
 (TM)

Release 3.1A John F. Collins, BioComputing Research Unit,
 Copyright (c) 1993-1995 University of Edinburgh, U.K.
 Distribution Rights by Oxford Molecular Ltd

Misfold - protein database search, using Smith-Waterman algorithm
 Run on: Mon Jun 19 15:56:49 2000; Master time 7.87 seconds
 518,264 Million cell updates/sec
 familiar output not generated.

Title: US-09-142-524A-3
 Description: (1-134) from US09142524A.gp
 Perfect Score: 1026
 Sequence: 1 MKVIVAFNFGPNRRVF1KRKTPRYDGIIAAYNPASKK 134

Scoring table: BMS 150
 Gap 11

Searched: 83857 seqs, 30454973 residues

Post-processing: Minimum Match 0%
 Testing first 45 summaries

Database: Swiss-Prot38
 Traversprot

Statistics: Mean 44.072; Variance 75.368; scale 0.555

Proj. No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed.
 and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | CS | ID | Description | Proj. No. |
|------------|-------|-------------|--------|----|-------------|------------------------------|-----------|
| 1 | 426 | 26.8 | 374 | 1 | SBP_CRYJA | SUGI BASIC PROTEIN PRECURSOR | 3,590-38 |
| 2 | 422 | 26.6 | 411 | 1 | MPA2_LACCA | PROBABILE POLYMERASE | 7,510-14 |
| 3 | 422 | 26.6 | 374 | 1 | PEL1_LACCA | PECTATE LYASE PRECURSOR | 3,540-10 |
| 4 | 413 | 25.9 | 434 | 1 | PEL1_LACCA | PECTATE LYASE PRECURSOR | 1,120-08 |
| 5 | 411 | 25.7 | 396 | 1 | MP11_LAMBAR | POLLEN ALLERGEN AMB A | 2,490-08 |
| 6 | 411 | 25.7 | 397 | 1 | MP11_LAMBAR | POLLEN ALLERGEN AMB A | 2,490-08 |
| 7 | 139 | 13.5 | 392 | 1 | MP4_LAMBAR | POLLEN ALLERGEN AMB A | 6,300-08 |
| 8 | 139 | 13.5 | 398 | 1 | PE56_LYCES | PROBABILE PECTATE LYASE | 9,300-08 |
| 9 | 137 | 13.4 | 398 | 1 | MP12_LAMBAR | POLLEN ALLERGEN AMB A | 1,370-07 |
| 10 | 137 | 13.4 | 404 | 1 | PE12_LYCES | STYLE DEVELOPMENT-SPEC | 1,370-07 |
| 11 | 131 | 12.6 | 397 | 1 | MPA2_LAMBAR | POLLEN ALLERGEN AMB A | 1,370-06 |
| 12 | 124 | 12.1 | 449 | 1 | PE59_LYCES | PROBABILE PECTATE LYASE | 1,850-05 |
| 13 | 105 | 10.2 | 110 | 1 | Y12K_SMSV4 | POSSIBLE 12 KD NUCLEIC | 1,590-02 |
| 14 | 100 | 9.7 | 3587 | 1 | TYCB_BACBR | TYROCIDINE SYNTHETASE | 8,180-02 |
| 15 | 99 | 9.6 | 2149 | 1 | RRPL_RVEV2 | RNA-DIRECTED RNA POLYM | 1,160-01 |
| 16 | 95 | 9.3 | 1070 | 1 | YHV4_YEAST | HYPOTHETICAL 123.0 KD | 4,180-01 |
| 17 | 93 | 9.1 | 539 | 1 | CH61_MCTCT | 60 KD CHAPERONIN 1 (PR | 7,830-01 |
| 18 | 89 | 8.9 | 638 | 1 | 60IM_COXBC | 60 KD INNER-MEMBRANE P | 1,450-00 |
| 19 | 90 | 8.6 | 339 | 1 | DPPL_ECOLI | DIPEPTIDE TRANSPORT SY | 1,970-00 |
| 20 | 90 | 8.6 | 377 | 1 | DDL_LEUME | D-ALANINE-D-ALANINE L | 1,970-00 |
| 21 | 90 | 8.5 | 450 | 1 | CPCL_RABIT | CYCLOCHROME P450 2C1 (E | 1,970-00 |
| 22 | 90 | 8.5 | 1074 | 1 | CASP_RAT | EXTRACELLULAR CALCIUM | 1,970-00 |
| 23 | 90 | 8.5 | 1085 | 1 | CASP_BOVIN | EXTRACELLULAR CALCIUM | 1,970-00 |

| | | | | | | | |
|----|----|-----|------|---|------------|-------------------------|----------|
| 24 | 89 | 8.7 | 309 | 1 | Y40Q_EH1SN | PROBABILE ABC TRANSP | 2,540-00 |
| 25 | 89 | 8.7 | 340 | 1 | G3P2_BACSO | GLYCERALDEHYDE 3-PH SP | 2,540-00 |
| 26 | 89 | 8.7 | 397 | 1 | PAPS_BACSO | POLY(A) POLYMERASE | 2,540-00 |
| 27 | 89 | 9.7 | 1059 | 1 | CERG_RAT | CERULOPLASMIN PRECUR | 2,540-00 |
| 28 | 89 | 9.7 | 1078 | 1 | CASP_HUMAN | EXTRACELLULAR CALCIUM | 2,540-00 |
| 29 | 88 | 8.6 | 293 | 1 | YF01_MYCTU | HYPOTHETICAL 33.0 KB P | 4,590-01 |
| 30 | 88 | 8.6 | 308 | 1 | RUEP_BACIN | RIBOFLAVIN BIOSYNTHES | 4,590-01 |
| 31 | 88 | 8.6 | 490 | 1 | CPEK_BACIN | CYCLOCHROME P450 2C1 (E | 1,590-01 |
| 32 | 87 | 8.5 | 328 | 1 | YXEL_BACSO | HYPOTHETICAL 37.2 KB P | 4,840-00 |
| 33 | 87 | 8.5 | 458 | 1 | CH14_YEAST | POTATIVE CELL SECRETAT | 4,840-00 |
| 34 | 87 | 8.5 | 608 | 1 | HYCO_ECOLI | FORMATE HYDROGENLYASE | 4,840-00 |
| 35 | 87 | 8.5 | 629 | 1 | FRET_YEAST | FERRIC REDUCTASE TRANS | 4,840-00 |
| 36 | 87 | 8.5 | 569 | 1 | SYVM_NEUCR | TYROSYL-TRNA SYNTHETAS | 4,840-00 |
| 37 | 86 | 8.4 | 381 | 1 | RGAL_HELPI | ROC SHAPE-DETERMINING | 5,470-00 |
| 38 | 86 | 8.4 | 419 | 1 | FAAA_HUMAN | FUMARYLACTATEASE (A | 5,470-00 |
| 39 | 86 | 8.4 | 461 | 1 | TRU_NEUCR | TUBULIN GAMMA CHAIN | 5,470-00 |
| 40 | 86 | 8.4 | 755 | 1 | COMP_RAT | CARLILASE OLIGOMERIN | 5,470-00 |
| 41 | 86 | 8.4 | 1062 | 1 | CERG_MOUSE | CERULOPLASMIN PRECURS | 5,470-00 |
| 42 | 85 | 8.3 | 262 | 1 | SR32_CAEEL | SRG-2 PROTEIN | 5,470-00 |
| 43 | 85 | 8.3 | 490 | 1 | CPCH_HUMAN | CYCLOCHROME P450 2C1 (E | 4,590-01 |
| 44 | 85 | 8.3 | 490 | 1 | CPHE_BABIT | CYCLOCHROME P450 2C1 (E | 4,590-01 |
| 45 | 85 | 8.3 | 5486 | 1 | TYOC_BACBR | TYROCIDINE SYNTHETASE | 8,180-02 |

ALIGNMENTS

| ID | SBP_CRYJA | STANDARD | PRT | 374 AA |
|----|---|----------|-----|--------|
| AC | P18532 | | | |
| DT | 01-NOV-1990 (Rel. 15, Created) | | | |
| DT | 01-NOV-1995 (Rel. 32, Last sequence update) | | | |
| DT | 15-DEC-1998 (Rel. 37, Last annotation update) | | | |
| DE | SUGI BASIC PROTEIN PRECURSOR (SBP) (MAJOR ALLERGEN CRV 1.1) (Crv 3.1) | | | |
| OS | Cryptomeria japonica (Japanese cedar) | | | |
| OC | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta | | | |
| OC | Euphyllophytes; Spermatophyta; Gniferales; Gniferales; Gniferales | | | |
| OC | Taxodiaceae; Cryptomeria | | | |
| RN | 1 | | | |
| RP | SEQUENCE FROM N.A. AND PARTIAL SEQUENCE | | | |
| RC | TISSUE=POLLEN | | | |
| RX | MEDLINE: 94183234 | | | |
| RA | Sone T., Komiyama N., Shimizu K., Kusakabe J., Morikawa K., | | | |
| RA | Kino K., | | | |
| RT | Cloning and sequencing of cDNA coding for Crv 1.1, a major allergen | | | |
| RT | of Japanese cedar pollen. | | | |
| RL | Biochem. Biophys. Res. Commun. 149:15-625(1994). | | | |
| RN | 12 | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | TISSUE=POLLEN | | | |
| RA | Nakada M., Kutsuno M., Tsuboi K., Katsuda S., Katsuda M. | | | |
| RA | Submitted (1994-1994) to the DNA Data Bank of Japan | | | |
| RN | 13 | | | |
| RP | SEQUENCE OF 22-41 | | | |
| RC | TISSUE=POLLEN | | | |
| RX | MEDLINE: 89031257 | | | |
| RA | Tachai M., Ando S., Usui M., Kurioto M., Sakaguchi M., Irouye S., | | | |
| RA | Matsuhashi T. | | | |
| RT | N-terminal amino acid sequence of a major allergen of Japanese Cedar | | | |
| RT | Pollen (Crv 1.1). | | | |
| RL | FEBS Lett. 239:329-332(1988). | | | |
| RN | 14 | | | |
| RP | CARBOHYDRATES | | | |
| RX | MEDLINE: 95003748 | | | |
| RA | Hajikata A., Matsumoto T., Kojima K., Ogawa H., | | | |
| RT | Antigenicity of the oligosaccharide moiety of the Japanese cedar | | | |
| RT | (Cryptomeria japonica) pollen allergen, Crv 1.1. | | | |
| RL | Int. Arch. Allergy Immunol. 105:198-202(1994). | | | |
| CC | 1-1 P.T.M. CONTAINS FUCCSE/XYLOSE-CONTAINING N-LINKED OLIGOSACCHARIDES. | | | |
| CC | 1-1 DISEASE: THIS PROTEIN IS THE MAJOR ALLERGEN OF JAPANESE CEDAR | | | |
| CC | POLLEN. THE MOST COMMON POLLEN ALLERGEN IN JAPAN | | | |
| CC | 1-1 MISCELLANEOUS: THE SEQUENCE OF CRV 1.1 FORM A IS SHOWN HERE. FORM | | | |
| CC | B DIFFERS IN SIX POSITIONS. | | | |
| CC | 1-1 SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1. | | | |

CC AT THEIR NON-REDUCING ENDS.
CC (-) DEVELOPMENTAL STAGE: MAXIMALLY EXPRESSED LATE IN POLLEN
CC DEVELOPMENT.
CC (-) SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X67158; CAA47630.1;
CC EMBL: X67159; CAA47631.1;
CC EMBL: X67160; CAA47632.1;
CC PIR: S26211; S26211;
CC PFAM: PF00544; pec_lyase.1;
CC PRINTS: PR00807; AMBALLERG
CC KW Lyase; Signal: 1 25 POTENTIAL;
CC FT SIGNAL: 26 397 PECTATE LYASE;
CC FT CHAIN: 26 397 POTENTIAL;
CC FT ACT_SITE: 272 272 POTENTIAL;
CC FT CARBOHYD: 134 134 POTENTIAL;
CC FT CARBOHYD: 227 227 POTENTIAL;
CC FT CONFLICT: 156 156 S -> C (IN MRNA);
CC FT CONFLICT: 189 190 GS -> SG (IN MRNA);
CC FT CONFLICT: 200 200 S -> D (IN MRNA);
CC FT CONFLICT: 202 202 H -> R (IN MRNA);
CC FT CONFLICT: 249 249 H -> N (IN MRNA);
CC SEQUENCE 367 AA: 4435; MW: 5048255DA7643F CRC64;
CC
CC Query Match 14.8%; Score 152; DB 1; Length 397;
CC Best Local Similarity 35.7%; Pred. No. 3.64e-10;
CC Matches 22; Conservative 15; Mismatches 22; Indels 1; Gaps 1;
CC
CC 125 YGVVCKENLNIIFKNKIKLSRELIVTSKRTIDGRFNVHONGAGIK-IQASNIILIS 153
CC 67 ITLKGRRPLWIFSGNKLKMKMYIAGYKTFDGRRAEVSYYVHVNGAKFIRRVQGIILAA 126
CC
CC RESULT 4
CC ID PELLILLO STANDARD: 981; 434 AA.
CC AC 146978
CC DT 01-FEB-1995 (Rel. 31; Created)
CC DT 01-FEB-1995 (Rel. 31; Last sequence update)
CC DT 15-DEC-1998 (Rel. 37; Last annotation update)
CC FT PECTATE LYASE PRECURSOR (EC 4.2.2.2).
CC FT Liliaceae, Viridiplantae (Stumpet 111).
CC CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC CC Euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
CC CC core eudicots; Asterales; euasterids I; Asteraceae; Asteroideae;
CC CC Liliaceae; Liliaceae.
CC RN
CC SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
CC TISSUE: POLLEN;
CC RA Refnar T., Griffith I.J., Kuo X.-C., Bond J.F., Rogers B.,
CC RA Klapper D.G.,
CC RT "Cloning of Amb 4 (antigen E), the major allergen family 1 sensit-
CC RT reased pollen."
CC RL J. Biol. Chem. 266:1229-1236(1991).
CC RN
CC SEQUENCE FROM N.A. AND VARIANTS
CC TISSUE: POLLEN;
CC RX MEDLINE: 92234570.
CC RA Griffith I.J., Pollock J., Klapper D.G., Rogers B., Bond J.F.,
CC RT "Sequence polymorphism of Amb 4 and Amb 5 in the major pollen aller-
CC RT gen Ambrosia alternifolia (Stumpet 111)."
CC RL Int. Arch. Allergy Appl. Immunol. 94:224-234(1994).
CC CC
CC (-) SUBMITTER: MONROE.
CC CC (-) TISSUE SPECIFICITY: POLLEN AND FLOWERS
CC CC (-) DISEASE: ONE OF THE MAJOR ALLERGENS OF THE PARNETED POLLEN.
CC CC (-) SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
CC CC AMB 4/AMB 4 II/CRY 3 I SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M63116; NOT-ANNOTATED_CDS.
CC EMBL: M80558; AAA32655.1;
CC PIR: A39099; A39099;
CC PIR: A53240; A53240;
CC PFAM: PF00544; pec_lyase.1;
CC PRINTS: PR00807; AMBALLERG
CC KW Antigen; Allergen; Signal: Multigene family; Polymorphism.

DR PFAM: PF00544; pec_lyase;
DR PRINTS: PR00807; AMBALLERG
DR KW Lyase; Signal: 1 22 POTENTIAL;
DR FT SIGNAL: 23 434 PECTATE LYASE;
DR FT CHAIN: 23 434 POTENTIAL;
DR FT ACT_SITE: 312 312 POTENTIAL;
DR FT CARBOHYD: 68 68 POTENTIAL;
DR FT CARBOHYD: 97 97 POTENTIAL;
DR SEQUENCE 434 AA: 48457 MW: 51936304H80C64 CRC64;
CC
CC Query Match 13.9%; Score 143; DB 1; Length 444;
CC Best Local Similarity 38.5%; Pred. No. 1.32e-08;
CC Matches 25; Conservative 14; Mismatches 22; Indels 4; Gaps 4;
CC
CC 140 FGLUR-MGVIOQRPLWIFGKSMVIRLKGRLINNDKIDGEGANVQ-I-ANAGNIIVF 176
CC 61 FAKLTGFLMGRPLWIFSG IRLKMYIAGYKTFDGRRAEVSYYVHVNGAKFIRRV 179
CC
CC 197 VHVVI 201
CC 120 VDGII 124
CC
CC RESULT 5
CC ID MP11:AMBAR STANDARD: 981; 396 AA.
CC AC P27759;
CC DT 01-AUG-1992 (Rel. 23; Created)
CC DT 01-AUG-1992 (Rel. 23; Last sequence update)
CC DT 15-DEC-1998 (Rel. 37; Last annotation update)
CC DE POLLEN ALLERGEN AMB A 1.1 PRECURSOR (ANTIGEN E) (ANTIGEN AMB A 1).
CC OS Ambrosia artemisiifolia (Stor. aqwee).
CC CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC CC Euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
CC CC core eudicots; Asterales; euasterids I; Asteraceae; Asteroideae;
CC CC Asteroideae; Heliantheae; Ambrosia.
CC RN
CC SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
CC TISSUE: POLLEN;
CC RX MEDLINE: 91093235.
CC RA Refnar T., Griffith I.J., Kuo X.-C., Bond J.F., Rogers B.,
CC RA Klapper D.G.,
CC RT "Cloning of Amb 4 (antigen E), the major allergen family 1 sensit-
CC RT reased pollen."
CC RL J. Biol. Chem. 266:1229-1236(1991).
CC RN
CC SEQUENCE FROM N.A. AND VARIANTS
CC TISSUE: POLLEN;
CC RX MEDLINE: 92234570.
CC RA Griffith I.J., Pollock J., Klapper D.G., Rogers B., Bond J.F.,
CC RT "Sequence polymorphism of Amb 4 and Amb 5 in the major pollen aller-
CC RT gen Ambrosia alternifolia (Stumpet 111)."
CC RL Int. Arch. Allergy Appl. Immunol. 94:224-234(1994).
CC CC
CC (-) SUBMITTER: MONROE.
CC CC (-) TISSUE SPECIFICITY: POLLEN AND FLOWERS
CC CC (-) DISEASE: ONE OF THE MAJOR ALLERGENS OF THE PARNETED POLLEN.
CC CC (-) SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
CC CC AMB 4/AMB 4 II/CRY 3 I SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M63116; NOT-ANNOTATED_CDS.
CC EMBL: M80558; AAA32655.1;
CC PIR: A39099; A39099;
CC PIR: A53240; A53240;
CC PFAM: PF00544; pec_lyase.1;
CC PRINTS: PR00807; AMBALLERG
CC KW Antigen; Allergen; Signal: Multigene family; Polymorphism.

FT SIGNAL 1 25 POTENTIAL.
 FT CHAIN 26 396 POLLEN ALLERGEN AMB A 1.1.
 FT MOD.RES 226 226 BLOCKED.
 FT VARIANT 92 92 E -> D.
 SQ SEQUENCE 396 AA: CCE75DE7B288841D CRC64:

Query Match 13.7% Score 141: DB 1: Length 396:
 Best Local Similarity 43.2% Pred. No. 2.89e-06:
 Matches 19: Conservative 8: Mismatches 17: Indels 0: Gaps 0:

DB 107 PAAAGNRRLWIFKNDKVMVNSUKITDGRGKVE1150
 QY 67 FTLMGRRLWIFSGNNKIKLMNYIAIKITDGRRAEVSVMVNSAKITPAAV1150

RESULT 1
 ID MP14_AMBAR STANDARD: PR1: 397 AA.

AC P2751:
 DT 01-AUG-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 25, Last sequence update)
 DT 15-DEC-1996 (Rel. 37, Last annotation update)
 DE POLLEN ALLERGEN AMB A 1.3 PRECURSOR (ANTIGEN E) (ANTIGEN AMB A 1.3)
 OS Ambrosia artemisiifolia (Short ragweed).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 OC core eudicots; Asterales; euasterids II; Asteraceae;
 OC Asteroideae; Heliantheae; Ambrosia.

RN [1]
 RP SEQUENCE FROM N.A. AND VARIANTS.
 RC TISSUE: POLLEN.
 RX MEDLINE: 91093245.
 RA Raftar T., Griffith I.J., Kio M.C., Bond J.F., Rogers B.L.,
 RA Klapper D.G., Pollock J., Klapper D.G., Rogers B.L., Nault A.K.,
 RT "Sequence polymorphism of Amb A 1 and Amb A 1.1: the major allergens
 RT in Ambrosia artemisiifolia (short ragweed)." J. Biol. Chem. 264:1229-1236(1991).

RN [2]
 RP SEQUENCE FROM N.A. AND VARIANTS.
 RC TISSUE: POLLEN.
 RX MEDLINE: 9223457.
 RA Griffith I.J., Pollock J., Klapper D.G., Rogers B.L., Nault A.K.,
 RT "Sequence polymorphism of Amb A 1 and Amb A 1.1: the major allergens
 RT in Ambrosia artemisiifolia (short ragweed)." J. Biol. Chem. 264:1229-1236(1991).

CC -1- SUBUNIT: MONOMER.
 CC -1- TISSUE SPECIFICITY: POLLEN AND FLOWERS.
 CC -1- DISEASE: ONE OF THE MAJOR ALLERGENS OF THE RAGWEED POLLEN.
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC AMB A 1/AMB A 1.1/CRY J 1 SUBFAMILY

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M62961: AAA32659.1: -
 CC PIR: M6C560: AAA32659.1: ALT_SEQ.
 CC PIR: C19099: C19099.
 CC PFAM: PF00544: pec_lyase: 1.
 CC PRINTS: PR0807: AMBALLERGEN.
 CC ACTIGN: Allergen: Signal: Multigene family: Polymorphism.
 FT SIGNAL 1 25 POTENTIAL.
 FT CHAIN 26 397 POLLEN ALLERGEN AMB A 1.3.
 FT MOD.RES 226 226 BLOCKED.
 FT VARIANT 48 48 L -> Y.
 SQ SEQUENCE 397 AA: C8DB41257590DD0A CRC64:

Query Match 13.7% Score 141: DB 1: Length 397:
 Best Local Similarity 37.9% Pred. No. 2.89e-06:
 Matches 22: Conservative 12: Mismatches 22: Indels 2: Gaps 2:

DB 106 PAAAGNRRLWIFKNDKVMVNSUKITDGRGKVE1150
 QY 67 FTLMGRRLWIFSGNNKIKLMNYIAIKITDGRRAEVSVMVNSAKITPAAV1150

RESULT 7
 ID MP14_AMBAR STANDARD: PR1: 392 AA.

AC P29744:
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 15-DEC-1996 (Rel. 37, Last annotation update)
 DE POLLEN ALLERGEN AMB A 1.4 PRECURSOR (ANTIGEN E) (ANTIGEN AMB A 1.3)
 OS Ambrosia artemisiifolia (Short ragweed).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 OC core eudicots; Asterales; euasterids II; Asteraceae;
 OC Asteroideae; Heliantheae; Ambrosia.

RN [1]
 RP SEQUENCE FROM N.A. AND VARIANTS.
 RC TISSUE: POLLEN.
 RX MEDLINE: 92234570.
 RA Griffith I.J., Pollock J., Klapper D.G., Rogers B.L., Nault A.K.,
 RT "Sequence polymorphism of Amb A 1 and Amb A 1.1: the major allergens
 RT in Ambrosia artemisiifolia (short ragweed)." J. Biol. Chem. 264:1229-1236(1991).

RN [2]
 RP SEQUENCE FROM N.A. AND VARIANTS.
 RC TISSUE: POLLEN.
 RX MEDLINE: 92234570.
 RA Griffith I.J., Pollock J., Klapper D.G., Rogers B.L., Nault A.K.,
 RT "Sequence polymorphism of Amb A 1 and Amb A 1.1: the major allergens
 RT in Ambrosia artemisiifolia (short ragweed)." J. Biol. Chem. 264:1229-1236(1991).

CC -1- SUBUNIT: MONOMER.
 CC -1- TISSUE SPECIFICITY: POLLEN AND FLOWERS.
 CC -1- DISEASE: ONE OF THE MAJOR ALLERGENS OF THE RAGWEED POLLEN.
 CC -1- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
 CC AMB A 1/AMB A 1.1/CRY J 1 SUBFAMILY

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M80562: AAA32670.1: -
 CC PIR: D51240: pec_lyase: 1.
 CC PRINTS: PR0807: AMBALLERGEN.
 CC ACTIGN: Allergen: Signal: Multigene family: Polymorphism.
 FT SIGNAL 1 25 POTENTIAL.
 FT CHAIN 26 392 POLLEN ALLERGEN AMB A 1.4.
 FT MOD.RES 226 226 BLOCKED.
 FT VARIANT 182 186 SHGPPV -> CNDLPFA
 SQ SEQUENCE 392 AA: 42542 MW: 40230.27 PPM: 0.0000000000000000

Query Match 13.5% Score 139: DB 1: Length 392:
 Best Local Similarity 15.4% Pred. No. 4.0e-05:
 Matches 20: Conservative 11: Mismatches 23: Indels 1: Gaps 1:

DB 106 PAAAGNRRLWIFKNDKVMVNSUKITDGRGKVE1150
 QY 67 FTLMGRRLWIFSGNNKIKLMNYIAIKITDGRRAEVSVMVNSAKITPAAV1150

RESULT 8
 ID PE56_LYCES STANDARD: PR1: 396 AA.

AC P15721:
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE PROBABLE PECTATE LYASE P56 /CURSOR (EC 4.2.2.2).
 GN LAT56.
 OS Lycopersicon esculentum (Tomato).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 OC core eudicots; Asterales; euasterids I; Solanales; Solanaceae;
 OC Solanum.

```

PP SEQUENCE FROM N.A.
RC STRAIN-CV: VF36; TISSUE=ANTHER;
RX MEDLINE: 9122485.
RA Wang R.A., Yarauchi J., Larabell S.K., Usin V.M., McCormick S.,
RI "Molecular and genetic characterization of two pollen-expressed genes
RT that have sequence similarity to pectate lyases of the plant pathogen
RI Erwinia."
RL Plant Mol. Biol. 14:17-28(1990).
RN [2]
RV REVISIONS.
RC STRAIN-CV: VF36; TISSUE=ANTHER;
RX Wang R.A.;
RI Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases
RI FUNCTION: MIGHT BE NEEDED DURING POLLEN DEVELOPMENT AND TUBE
CC GROWTH.
CC CATALYTIC ACTIVITY: ELIMINATIVE CLEAVAGE OF PECTATE TO GIVE
CC OLIGOSACCHARIDES WITH 4-DEOXY-ALPHA-D-MANN-4-ENURONOSYL GROUPS
CC AT THEIR NON-REDUCING ENDS.
CC TISSUE SPECIFICITY: EXPRESSED IN ANTHERS AND POLLEN.
CC SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X15500; CAA33524.1;
CC PIR: S08383; SC8383;
CC PFAM: PF00544; pect_lyase; 1;
CC PRINTS: PR00807; AMBALERGEN.
CC Lyase: Multimeric family: Signal.
CC SIGNAL: 1 27 GR 22 (POTENTIAL);
CC CHAIN: 28 398 PROBABLE PECTATE LYASE P56.
CC ACT_SITE: 273 273 POTENTIAL.
CC CARBOHYD: 135 135 POTENTIAL.
CC CARBOHYD: 228 228 POTENTIAL.
CC SEQUENCE 398 AA; 44563 MW; 806762505; iBC708 CRC64;
CC
CC Query Match 13.5%; Score 138; DB 1; Length 398;
CC Best Local Similarity 35.5%; Pred. No. 9,30e-08;
CC Matches 22; Conservative 17; Mismatches 5; Index 5; Gaps 4;
CC
CC 106 FGVIKQPKWITFARSMR:RLTRELIVSSNKTIQGR-GK--YVHANGAGIKISASNV: 162
CC 113 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
CC 67 FILMRRPLWILFSGNNMNIKMKMVMYAGYKTFDGRRAEVSVHV-NGAKF-IRRVGGII: 124
CC
CC 163 15 164
CC 125 AA 126
CC
CC RESULT 9
CC ID WP12 AMBAR STANDARD: PRI: 398 AA.
CC AC P27760;
CC DT 01-AUG-1992 (Rel. 23, Created)
CC DT 01-JUL-1992 (Rel. 23, Last sequence update)
CC DT 15-DEC-1998 (Rel. 37, Last annotation update)
CC DE POLLEN ALLERGEN AMB A 1.2 PRECURSOR (ANTIGEN E)
CC OS Ambrosia artemisiifolia (Short ragweed).
CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
CC core eudicots; Asteridae; euasterids II; Asterales; Asteraceae;
CC Asteroideae; Heliantheae; Ambrosia.
CC [-]
CC RN SEQUENCE FROM N.A.
CC RC TISSUE=POLLEN;
CC RX MEDLINE: 91093235.
CC RA Rafnar T., Griffith I.J., Kuo M.-C., Bond J.F., Rogers B.L.,
CC Klapper D.G.;

```

```

RT "Cloning of Amb a 1 (antigen E), the major allergen family of short
ragweed pollen."
RL J. Biol. Chem. 266:1229-1236(1991).
RN [2]
RV SEQUENCE FROM N.A. AND VARIANTS
RC TISSUE=POLLEN;
RX MEDLINE: 92234570.
RA Griffith I.J., Pollack J., Kupper D.G., Rogers B.L., Nault A.K.;
RI "Sequence polymorphism of Amb a 1 and Amb a 11, the major allergens
in Ambrosia artemisiifolia (Short ragweed)."
RL Int. Arch. Allergy Appl. Imm Biol. 96:296-304(1991).
RN [2]
RV SUBUNIT: MONOMER.
CC TISSUE SPECIFICITY: POLLEN AND FLOWERS.
CC DISEASE: ONE OF THE MAJOR ALLERGENS OF THE RAGWEED POLLEN.
CC SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1.
CC AMB A 1/AMB A 11/CRY J. 1 SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed, usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: M52981; AAA32656.1;
CC EMBL: M80559; AAA32667.1;
CC PIR: B39099; B39099;
CC PFAM: PF00544; pect_lyase; 1;
CC PRINTS: PR00807; AMBALERGEN.
CC Antigen: Allergen: Signal; Multimeric family: Polymorphic.
CC SIGNAL: 1 25 POTENTIAL.
CC CHAIN: 26 398 POLLEN ALLERGEN AMB A 1.2.
CC MOD_RES: 226 226 BLOCKED.
CC VARIANT: 345 345 N -> K.
CC VARIANT: 381 381 N -> I.
CC SEQUENCE 398 AA; 43664 MW; 020066209B74150 CRC54;
CC
CC Query Match 13.4%; Score 117; DB 1; Length 398;
CC Best Local Similarity 39.7%; Pred. No. 1.37e-07;
CC Matches 23; Conservative 10; Mismatches 23; Index 2; Gaps 2;
CC
CC 109 FFAAQNRPILKIFKNNVHNLNDELAVNSKTIQGRGVKVNVA-GLILNN-VKSI: 164
CC 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
CC 67 FILMRRPLWILFSGNNMNIKMKMVMYAGYKTFDGRRAEVSVHV-NGAKF-IRRVGGII: 124
CC
CC RESULT 10
CC ID 9612 LYCES STANDARD: PRI: 474 AA.
CC AC P24396;
CC DT 01-MAR-1992 (Rel. 21, Created)
CC DT 01-MAR-1992 (Rel. 21, Last sequence update)
CC DT 15-JUL-1999 (Rel. 38, Last annotation update)
CC DE STYLE DEVELOPMENT-SPECIFIC PROTEIN 9612 PRECURSOR.
CC GN 9612.
CC OS Lycopersicon esculentum (Tomato).
CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
CC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
CC core eudicots; Asteridae; euasterids I; Solanales; Solanaceae;
CC Solanum.
CC [-]
CC RN SEQUENCE FROM N.A.
CC RC STRAIN-CV: VF36; TISSUE=PISTIL;
CC RX MEDLINE: 91117185.
CC RA Badellier K.A., Smith A.G., Gasser C.S.;
RI "Regulation of a stylar transmitting tissue-specific gene in
wild-type and transgenic tomato and tobacco."
RL Mol. Gen. Genet. 224:183-192(1990).
CC FUNCTION: MAY HAVE A ROLE IN THE DEVELOPMENT OF THE TRANSMITTING
CC TISSUE OF THE STYLE AND/OR IN THE EVENTS RELATED TO POLLINATION
CC SUCH AS SOME ASPECT IN THE FACILITATION OF COMPATIBLE POLLEN TUBE
CC GROWTH.
CC [-]
CC SUBCELLULAR LOCATION: SECRETED.

```


CC 1- TISSUE SPECIFICITY: PREDOMINANTLY FOUND IN THE PISTIL WHERE IT IS
CC FOUND IN THE OUTER FIVE LAYERS OF THE STRANDS OF TRANSMITTING
CC TISSUE WITHIN THE UPPER TWO-THIRDS OF THE STYLE. FOUND AT MUCH
CC LOWER LEVELS IN THE ANthers AND VEGETATIVE ORGANS.
CC 2- DEVELOPMENTAL STAGE: MAXIMUM LEVELS ARE FOUND DURING ANTHESIS.
CC 3- SIMILARITY: 54% IDENTICAL TO TOMATO PROTEINS P59 (AC P.5722)
CC AND P56 (AC P.5722).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

EMBL: X55193; CAA36979.1;
PIR: S12209; S12209;
PFAM: PF00544; PECT_LYASE; 1;
PRINTS: PR06807; AMBALLERGEN;
KW Signal;
FT SIGNAL 1 20 POTENTIAL;
FT CHAIN 21 404 STYLE DEVELOPMENT-SPECIFIC PROTEIN 9612;
FT CARBOHYD 37 37 POTENTIAL;
FT CARBOHYD 191 191 POTENTIAL;
SQ SEQUENCE 424 AA: 44298 MW: 82656.95; 2688675 CRG64;

Query Match 22.48; Score 137; DB 1; Length: 404;
Best Local Similarity 54.5%; Pred. No. 1370-87;
Matches 18; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

DB 120 PRLWIFQRMVIOIKQELVMSNYKIFGRGASV 152
|||||
CY 74 PRLWIFQRMVIOIKQELVMSNYKIFGRGASV 152

RESULT 11
ID MP22-AMBAR STANDARD; PRT: 397 AA;
AC 927762;
DT 01-APR-1992 (Rel. 23, Created)
DI 01-APR-1992 (Rel. 23, Last sequence update)
UI 01-DEC-1998 (Rel. 37, Last annotation update)
DE POLLEN ALLERGEN AMB A 2 PRECURSOR (ANTH: K) (ANTH: AMB A 1);
OS Ambrosia artemisiifolia (short ragweed);
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC core eudicots; Asteridae; easterids I; Asterales; Asteroideae;
OC Asteroideae; Bellanthiales; Ambrosia;
RN 11;
RC TISSUE: FLOWER;
RX MEDLINE: 92234570;
RA Rogers B.L., Morgenstern J.P., Griffith I.J., Yu X.-B.,
RA Connell C.M., Brauer A.W., King T.P., Garman R.D., Kuo M.-C. C.;
FT "Complete sequence of the allergen Amb A 1, the major allergen
FT expression and reactivity with T cells from ragweed allergic
FT patients".
RT J. Immunol. 147:2547-2552(1991).
RL 12;
RN 12;
RP SEQUENCE FROM N.A., AND VARIANTS;
RC TISSUE: POLLEN;
RX MEDLINE: 92234570;
RA Griffith I.J., Pollock J., Kapper D.G., Rogers B.L., Nault A.K.,
RA "Sequence polymorphism of Amb A 1 and Amb A 2, the major allergens
RT in Ambrosia artemisiifolia (short ragweed)".
RL Int. Arch. Allergy Appl. Immunol. 96:296-304(1991).
CC 1- SUBUNIT: MONOMER;
CC 2- TISSUE SPECIFICITY: POLLEN AND FLOWERS;
CC 3- DISEASE: ONE OF THE MAJOR ALLERGENS OF THE RAGWEED POLLEN;
CC 4- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1;
CC AMB A 1/AMB A 2/CRY J SUBFAMILY;
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

EMBL: M80561; AAA32571.1;
PIR: A45459; A45459;
PFAM: PF00544; PECT_LYASE; 1;
PRINTS: PR06807; AMBALLERGEN;
KW Antigen; Allergen; Signal; Polysaccharide;
FT SIGNAL 25 POTENTIAL;
FT CHAIN 26 297 POLLEN ALLERGEN AMB A 2;
FT MOD_RES 725 725 BLUENES;
FT VARIANT 70 70 N - C (DETECTED ONLY IN FLOWER (NA));
FT VARIANT 138 138 K - T (DETECTED ONLY IN FLOWER (NA));
FT VARIANT 321 321 K - R (DETECTED ONLY IN FLOWER (NA));
SQ SEQUENCE 397 AA: 44082 MW: 68617.94; 101010 CRG64;

Query Match 12.88; Score 14; DB 1; Length: 473;
Best Local Similarity 42.3%; Pred. No. 1370-86;
Matches 22; Conservative 8; Mismatches 40; Indels 0; Gaps 2;

DB 114 RPLWIFQRMVIOIKQELVMSNYKIFGRGASV 152
|||||
CY 73 RPLWIFQRMVIOIKQELVMSNYKIFGRGASV 152

RESULT 12
ID P559-LYCES STANDARD; PRT: 445 AA;
AC P15722;
DT 01-APR-1990 (Rel. 14, Created)
DI 01-APR-1990 (Rel. 14, Last sequence update)
UI 15-JUL-1999 (Rel. 38, Last annotation update)
DE PROBABLE PECTATE LYASE P59 PRECURSOR (F 4.2.2.2)
OS Lycopersicon esculentum (Tomato)
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC euphyllophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC core eudicots; Asteridae; easterids I; Solanales; Solanales;
RN 1;
RC SEQUENCE FROM N.A.
RC SPRAIN-CV. VF35; TISSUE: ANTH: K;
RX MEDLINE: 92322485;
RA Wang P.A., Yamamoto K.;
FT "Molecular and genetic characterization of the gene encoding pectate lyase
FT that have sequence similarity to pectate lyase 1 from Arabidopsis
FT thaliana".
RL Plant Mol. Biol. 34:17-26(1990).
RN 1;
RC FUNCTION: MIGHT BE NEEDED DURING POLLEN DEVELOPMENT AND FERTILIZATION;
CC 1- CATALYTIC ACTIVITY: ELIMINATIVE CLEAVAGE OF PECTATE TO GIVE
CC OLIGOSACCHARIDES WITH 4-DEOXY-ALPHA-D-MANN-4-ENURON-SYL GROUPS
CC AT THEIR NON-REDUCING ENDS;
CC 2- TISSUE SPECIFICITY: EXPRESSED IN ANthers AND POLLEN;
CC 3- SIMILARITY: BELONGS TO THE POLYSACCHARIDE LYASE FAMILY 1;
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).

EMBL: X15499; CAA33523.1;
PIR: S27098; S27098;
PFAM: PF00544; PECT_LYASE; 1;
PRINTS: PR06807; AMBALLERGEN;
KW Lyase; Multigene family; Signal;
FT SIGNAL 1 22 POTENTIAL;


```

ID  RP2-RVEVZ      STANDARD:      PRI: 2149 AA:
AC  P27316
DI  01-AUG-1992 (rel. 23, Created:
DI  01-AUG-1992 (rel. 23, Last sequence update)
DI  01-AUG-1992 (rel. 23, Last annotation update)
DE  RNA-DIRECTED RNA POLYMERASE (EC 2.7.7.45) (L PROTEIN).
OR  Rift valley fever virus (strain ZH-548 M2) (RVFV).
OC  Viruses; ssRNA negative-strand viruses; Bunyaviridae; Phlebovirus.
RN  [1]
RP  SEQUENCE FROM N.A.
PX  MEDLINE: 92020236.
RA  Muller A., Argenzini G., Bouloy M., Prehaud C., Bishop D.H.L.
RI  "Completion of the genome sequence of Rift Valley fever phlebovirus
RI  indicates that the L RNA is negative sense and codes for a putative
RI  transcriptase-replicase."
RL  Nucleic Acids Res. 19:5433-5437(1991).
OC  -- CATALYTIC ACTIVITY: N NUCLEOSIDE TRIPHOSPHATE + N PYROPHOSPHATE
OC  + RNA(N).
OC  This SWISS-PROT entry is copyright. It is produced through a collaboration
OC  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
OC  the European Bioinformatics Institute. There are no restrictions on its
OC  use by non-profit institutions as long as its content is in no way
OC  modified and this statement is not removed. Usage by and for commercial
OC  entities requires a license agreement (see http://www.isb-sib.ch/annouce/
OC  or send an email to license@isb-sib.ch).
OC  -----
DR  EMBL: X56444; CAA39836.1; -.
OR  PIR: S19576; S19576
KW  Transcriptase; Nucleotidyltransferase; RNA-directed RNA polymerase.
SQ  SEQUENCE: 2149 AA; 243569 MW; 6D5739C679A58D7 CRC64;
Query Match: 9.6%; Score 99; DB 1; Length 2149;
Best Local Similarity 34.4%; Pred. No. 1,160-01;
Matches 22; Conservative 14; Mismatches 22; Indels 4; Gaps 4;
DE  988 TSDAPKNAAGHIVTKFALMCEFTSPKWWPL-ITGDSMTKKMMNNLN-YLKILDGH 1045
DY  1 111 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
DY  45 TISTGPKKNNPWLQQA-KLICEIMGRRLWITFGNNMNIKLKMPYIASY-KTIDGR 102
GN  1046 R 1045
QY  103 R 103
Search completed: Mon Jun 19 15:57:01 2000
JOL time : 12 secs.
```

M E M O R A N D U M

(TM)

Kellogg 3-11A John F. Collins, Biocomputing Research Unit,
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Meshch_pf protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Jun 13 15:57:18 2000; MasPar time 18.25 Seconds
Tabular output not generated. 509,207 Million cell updates/sec

File: US-09-142-524A-3
Description: (1-134) from US09142524A.pep
Perfect Score: 1026
Sequence: 1 MKVIVANQGFENRRVRIK.....KEIRRVGGIIIAAYQNPAWK 134

Scoring table: BAW 150
Car 11

Searched: 225878 seqs, 59334122 residues

Post-processing: Minimum Match 0%
Listing first 45 summaries

Database: sprentbl2
1:SP_dichaea 2:sp_bacteria 3:sp_fungi 4:sp_human
5:sp_invertebrate 6:sp_mammal 7:sp_mhc 8:sp_organella
9:sp_phase 10:sp_plant 11:sp_protein 12:sp_unclassified
13:sp_vertebrate 14:sp_virus

Statistics: Mean 42.950; Variance 75.878; scale 0.566

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

| Result | | Query | | Match | | Length | | EB | | ID | | Description | | Pred. No. | |
|--------|-----|-------|------|-------|--------|--------|--|----|--|----|--|-------------------------|--|-----------|--|
| 1 | 224 | 21.8 | 567 | 10 | Q92NU7 | | | | | | | POLLEN MAJOR ALLERGEN | | 3.59e-24 | |
| 2 | 210 | 20.5 | 355 | 10 | Q96385 | | | | | | | CHAOI PRECURSOR | | 2.24e-21 | |
| 3 | 150 | 14.6 | 191 | 10 | Q23666 | | | | | | | PUTATIVE PECTATE LYASE | | 5.53e-10 | |
| 4 | 148 | 14.4 | 431 | 10 | Q23017 | | | | | | | TIG11.7 PROTEIN | | 1.26e-09 | |
| 5 | 142 | 13.7 | 455 | 10 | Q64510 | | | | | | | PUTATIVE PECTATE LYASE | | 2.18e-08 | |
| 6 | 140 | 13.6 | 350 | 10 | Q65388 | | | | | | | FL2F1.22 PROTEIN | | 3.27e-08 | |
| 7 | 138 | 13.5 | 459 | 10 | Q23665 | | | | | | | PUTATIVE PECTATE LYASE | | 7.30e-08 | |
| 8 | 136 | 13.3 | 394 | 10 | Q65456 | | | | | | | PECTATE LYASE LIKE PRC | | 1.62e-07 | |
| 9 | 134 | 13.1 | 394 | 10 | Q65457 | | | | | | | PECTATE LYASE LIKE PRC | | 3.58e-07 | |
| 10 | 134 | 13.1 | 425 | 10 | Q24416 | | | | | | | PECTATE LYASE | | 3.58e-07 | |
| 11 | 133 | 13.0 | 226 | 10 | Q23667 | | | | | | | PUTATIVE PECTATE LYASE | | 5.12e-07 | |
| 12 | 131 | 12.8 | 438 | 10 | Q43852 | | | | | | | POLLEN SPECIFIC PECTAT | | 1.17e-06 | |
| 13 | 130 | 12.7 | 450 | 10 | Q40319 | | | | | | | PECTATE LYASE HOMOLOG | | 1.72e-06 | |
| 14 | 129 | 12.6 | 227 | 10 | Q23668 | | | | | | | PUTATIVE PECTATE LYASE | | 2.55e-06 | |
| 15 | 124 | 12.1 | 274 | 10 | Q24159 | | | | | | | PUTATIVE PECTATE LYASE | | 1.75e-05 | |
| 16 | 122 | 11.9 | 401 | 10 | Q24554 | | | | | | | PECTATE LYASE PRECURSO | | 3.74e-05 | |
| 17 | 118 | 11.5 | 399 | 10 | Q43783 | | | | | | | PECTATE LYASE (EC 4.2. | | 1.68e-04 | |
| 18 | 113 | 10.7 | 4450 | 10 | Q44922 | | | | | | | GRAM-ICIDIN'S SYNTHETAS | | 3.15e-03 | |
| 19 | 103 | 10.0 | 1484 | 10 | Q9X585 | | | | | | | POLYPROTEIN | | 3.72e-02 | |
| 20 | 100 | 9.7 | 238 | 10 | Q02901 | | | | | | | HYPOTHETICAL 27.5 KD P | | 1.04e-01 | |

| | | | | | | | | | | | | | | | |
|----|-----|-----|------|----|--------|--|--|--|--|--|--|------------------------|--|----------|--|
| 21 | 100 | 9.7 | 333 | 10 | OB1494 | | | | | | | F9D12.7 PROTEIN | | 1.04e-01 | |
| 22 | 98 | 9.6 | 626 | 5 | Q46039 | | | | | | | EG103B4.3 PROTEIN | | 2.04e-01 | |
| 23 | 96 | 9.4 | 190 | 2 | Q9X000 | | | | | | | HYPOTHETICAL 22.4 KD P | | 3.97e-01 | |
| 24 | 95 | 9.3 | 108 | 14 | Q09783 | | | | | | | CAPSIN | | 5.54e-01 | |
| 25 | 95 | 9.3 | 3229 | 5 | Q26912 | | | | | | | PROTEIN 1 OF A LIPID | | 6.56e-01 | |
| 26 | 94 | 9.2 | 554 | 2 | P26531 | | | | | | | HYPOTHETICAL 53.9 KD P | | 7.06e-01 | |
| 27 | 94 | 9.2 | 815 | 5 | Q23850 | | | | | | | VACUOLAR PROTON ATPASE | | 7.45e-01 | |
| 28 | 93 | 9.1 | 875 | 11 | P97475 | | | | | | | ALKALINE PHOSPHATASE | | 1.04e-00 | |
| 29 | 93 | 9.1 | 875 | 11 | P97476 | | | | | | | ALKALINE PHOSPHATASE | | 1.04e-00 | |
| 30 | 93 | 9.1 | 875 | 11 | C4490 | | | | | | | PR13-6 ANTIGEN | | 1.04e-00 | |
| 31 | 93 | 9.1 | 876 | 11 | P70541 | | | | | | | PHOSPHOGLYCERATE 1 | | 1.04e-00 | |
| 32 | 93 | 9.1 | 1347 | 2 | Q38426 | | | | | | | XILANASE | | 1.04e-00 | |
| 33 | 92 | 9.0 | 215 | 3 | Q95055 | | | | | | | A2 MATING TYPE 1 | | 1.46e-00 | |
| 34 | 92 | 9.0 | 360 | 2 | Q92359 | | | | | | | HYPOTHETICAL 41.9 KD P | | 1.46e-00 | |
| 35 | 91 | 8.9 | 215 | 2 | Q56807 | | | | | | | HYPOTHETICAL 25.5 KD P | | 2.03e-00 | |
| 36 | 91 | 8.9 | 426 | 10 | Q80447 | | | | | | | S GLYCOPROTEIN 1 | | 2.03e-00 | |
| 37 | 91 | 8.9 | 438 | 3 | Q13741 | | | | | | | POTATIVE RNA-BINDING P | | 2.03e-00 | |
| 38 | 91 | 8.9 | 596 | 10 | Q49550 | | | | | | | 1-PROSPHATIDYLIN-SITOL | | 2.03e-00 | |
| 39 | 91 | 8.9 | 652 | 14 | Q9MLD7 | | | | | | | ENVELOPE GLYCOPROTEIN | | 2.03e-00 | |
| 40 | 90 | 8.8 | 66 | 1 | Q29714 | | | | | | | PROTEIN TRANSFERASE S | | 2.76e-00 | |
| 41 | 90 | 8.8 | 146 | 11 | Q88982 | | | | | | | EXTRACELLULAR CALCIUM | | 2.76e-00 | |
| 42 | 90 | 8.8 | 377 | 2 | P71454 | | | | | | | D-ALA-D-ALA LIGASE | | 2.76e-00 | |
| 43 | 90 | 8.8 | 450 | 10 | Q40844 | | | | | | | VICILIN-LIKE STORAGE P | | 2.76e-00 | |
| 44 | 90 | 8.8 | 589 | 1 | Q33723 | | | | | | | DNA LIGASE | | 2.76e-00 | |
| 45 | 89 | 8.7 | 300 | 14 | Q65561 | | | | | | | ULF POLYPEPTIDE | | 3.75e-00 | |

ALIGNMENTS

RESULT 1
ID Q92NU7 PRELIMINARY: PRT: 367 AA.
AC Q92NU7:
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DI 01-NOV-1999 (Tremblrel. 12, Last annotation update)
DE POLLEN MAJOR ALLERGEN 1-2
OS Juniperus ashei (Orkard white cedar)
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta
OC euphyllophytes: Spermatophyta: Coniferopsida: Coniferales
OC Taxodiaceae: Juniperus
RN 11
RP SEQUENCE FROM N.A.
RA MIDORO-HOSIUTI T.M., GOLDBL. A.M., KOSKOV A., WOOD T.G.
RA BROOKS E.G.:
RT "Molecular cloning of mount cedar (Juniperus ashei) pollen major
allergen, Jun a 1.1."
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF106563; AAD03609.1
DR EMBL: AF106562; AAD03608.1
DR EMBL: J5544; J048128.1
DR MENDEL: J5545; J048128.1
DR MENDEL: J5545; J048128.1
SQ SEQUENCE 367 AA: 35924 MW: 422464 Da

Query Match 21.8% Score 224; EB 10; Length 567;
Best Local Similarity 55.6%; Pred. No. 3.59e-24;
Matches 29; Conservative 10; Mismatches 13; Indels 0; Gaps 0;
DB 87 KALWIFSONNKKLMPLVYAGHKITDGGADVHLGNGGQCLPFMKVSHVI 138
OY 73 RELWIFSONNKKLMPLVYAGHKITDGGADVHLGNGGQCLPFMKVSHVI 124
RESULT 2
ID Q96385 PRELIMINARY: PRT: 375 AA.
AC Q96385:
DT 01-FEB-1997 (Tremblrel. 02, Created)
DI 01-FEB-1997 (Tremblrel. 0, Last sequence update)
DT 01-NOV-1999 (Tremblrel. 12, Last annotation update)
DE CHAOI PRECURSOR
OS Chamaecyparis obtusa
OC Chamaecyparidaceae: Streptophyta: Embryophyta: Tracheophyta
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta
OC euphyllophytes: Spermatophyta: Coniferopsida: Coniferales
OC Taxodiaceae: Chamaecyparis

```

PN 11
PF SEQUENCE FROM N.A.
RC TISSUE: POLLEN
RX MEDLINE: 96265194
RA SUZUKI M., KONIYAKI N., ITOH H., SONE T., KUNE K., TAKAGI I.,
RA OHRA N.
RC "Purification, characterization and molecular cloning of Cha o 1, a
RC major allergen of Chara acutyrpis obtusa (Japanese rysses) pollen."
RC Mol. Immunol. 33:451-460(1996).
RC EMBL: D45404; BAA08245; J01000
RC MENDEL: 7626; Arabid1088; 7626
RC PFAM: PF00554; pec_lyase; 1
RC PRINTS: PR00807; AMBALLERGEN
KW Signal:
FT SIGNAL: 1 21 POTENTIAL
FT CHAIN: 22 375 CHAOL
SQ SEQUENCE 375 AA: 40259 MW: A0961492 CRC32:

Query Match 20.5% Score 210; DB 10; Length 375;
Best Local Similarity 57.7% Pred. No. 2,24e-21;
Matches 30; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

DB E7 SCWITFSGNNKIKMPHYIAGYKIDGPRAEVSVYVHVGAKFIRVDGII 124
QY 73 RPLWIFSGNNKIKMPHYIAGYKIDGPRAEVSVYVHVGAKFIRVDGII 124

RESULT 7
ID Q23666 PRELIMINARY: PRT: 181 AA.
AC Q23666 01-JAN-1998 (TrEMBLrel. 05; Created)
DI 01-AUG-1998 (TrEMBLrel. 05; Last sequence update)
DI 01-NOV-1999 (TrEMBLrel. 12; Last annotation update)
DE PUTATIVE PECTATE LYASE (FRAGMENT).
IN Arabidopsis thaliana (Mouse-ear cress).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC eukaryophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC core eudicots; Rosidae; eurosids 1; Brassicales; Brassicaceae;
OC Arabidopsis.
(1)
PN SEQUENCE FROM N.A.
RX MEDLINE: 97422403
RA KULAKOSKAS K., MCCORMICK S.
RC "Identification of the tobacco and Arabidopsis homologues of the
RC pollen-expressed LATE5 gene of tomato."
RC Plant Mol. Biol. 34:899-914(1997).
RC EMBL: U84620; AAB57663; J01000
RC MENDEL: 2556; Arabid1088; 2556
RC PFAM: PF00544; pec_lyase; 1
KW Cress.
FT SIGNAL: 1 181
FT CHAIN: 22 375
SQ SEQUENCE 375 AA: 27088 MW: EF342656 CRC32:

Query Match 14.6% Score 150; DB 10; Length 181;
Best Local Similarity 42.3% Pred. No. 5,53e-10;
Matches 22; Conservative 10; Mismatches 19; Indels 1; Gaps 1;

DB 53 RPLWIFSGNNKIKMPHYIAGYKIDGPRAEVSVYVHVGAKFIRVDGII 103
QY 73 RPLWIFSGNNKIKMPHYIAGYKIDGPRAEVSVYVHVGAKFIRVDGII 124

RESULT 4
ID Q23017 PRELIMINARY: PRT: 431 AA.
AC Q23017 01-JAN-1998 (TrEMBLrel. 05; Created)
DI 01-JAN-1998 (TrEMBLrel. 05; Last sequence update)
DI 01-NOV-1999 (TrEMBLrel. 12; Last annotation update)
DE TIG1.7 PROTEIN.
IN Arabidopsis thaliana (Mouse-ear cress).
OS

```

```

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC eukaryophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC core eudicots; Rosidae; eurosids 1; Brassicales; Brassicaceae;
OC Arabidopsis.
(1)
PN SEQUENCE FROM N.A.
RX MEDLINE: 96265194
RA SUZUKI M., KONIYAKI N., ITOH H., SONE T., KUNE K., TAKAGI I.,
RA OHRA N.
RC "Purification, characterization and molecular cloning of Cha o 1, a
RC major allergen of Chara acutyrpis obtusa (Japanese rysses) pollen."
RC Mol. Immunol. 33:451-460(1996).
RC EMBL: D45404; BAA08245; J01000
RC MENDEL: 7626; Arabid1088; 7626
RC PFAM: PF00554; pec_lyase; 1
RC PRINTS: PR00807; AMBALLERGEN
KW Signal:
FT SIGNAL: 1 21 POTENTIAL
FT CHAIN: 22 375 CHAOL
SQ SEQUENCE 375 AA: 40259 MW: A0961492 CRC32:

Query Match 14.4% Score 148; DB 10; Length 431;
Best Local Similarity 40.4% Pred. No. 1,24e-09;
Matches 22; Conservative 11; Mismatches 19; Indels 1; Gaps 1;

DB 146 RPLWIFSGNNKIKMPHYIAGYKIDGPRAEVSVYVHVGAKFIRVDGII 146
QY 73 RPLWIFSGNNKIKMPHYIAGYKIDGPRAEVSVYVHVGAKFIRVDGII 124

RESULT 5
ID Q4510 PRELIMINARY: PRT: 455 AA.
AC Q4510 01-AUG-1998 (TrEMBLrel. 07; Created)
DI 01-AUG-1998 (TrEMBLrel. 07; Last sequence update)
DI 01-NOV-1999 (TrEMBLrel. 12; Last annotation update)
DE PUTATIVE PECTATE LYASE.
IN Arabidopsis thaliana (Mouse-ear cress).
OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC eukaryophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
OC core eudicots; Rosidae; eurosids 1; Brassicales; Brassicaceae;
OC Arabidopsis.
(1)
PN SEQUENCE FROM N.A.
RX MEDLINE: 96265194
RA SUZUKI M., KONIYAKI N., ITOH H., SONE T., KUNE K., TAKAGI I.,
RA OHRA N.
RC "Purification, characterization and molecular cloning of Cha o 1, a
RC major allergen of Chara acutyrpis obtusa (Japanese rysses) pollen."
RC Mol. Immunol. 33:451-460(1996).
RC EMBL: D45404; BAA08245; J01000
RC MENDEL: 7626; Arabid1088; 7626
RC PFAM: PF00554; pec_lyase; 1
RC PRINTS: PR00807; AMBALLERGEN
KW Signal:
FT SIGNAL: 1 21 POTENTIAL
FT CHAIN: 22 375 CHAOL
SQ SEQUENCE 375 AA: 40259 MW: A0961492 CRC32:

Query Match 13.7% Score 141; DB 10; Length 455;
Best Local Similarity 45.1% Pred. No. 2,18e-08;
Matches 23; Conservative 11; Mismatches 15; Indels 4; Gaps 4;

```

Cy 74 PLWIFSGNNNIKMKMPMYIAGYKTFGGRFAEVSIVHVGAKF-IRRVGGI: 124

065456 PRELIMINARY: PRT: 394 AA.

063456: 01-AUG-1996 (TREMBlre: 07: created;
01-AUG-1998 (TREMBlre: 07: 1st sequence update)
01-NOV-1999 (TREMBlre: 12: 1st annotation update)
PECTATE LYASE LIKE PROTEIN.
FIN20.180.
Arabisopsis thaliana (Mouse-ear cress);
Eukaryota; Viridiplantae; Streptophyta; Erythropbyta; Arabidopsidaceae;
e-phylophyes; Spermatophyta; Magnoliophyta; eudicotyledons
core eudicots; Rosidae; eucotsids 1; Brassicales; Brassicaceae;
Arabidopsis.
[1]
SEQUENCE FROM N.A.
REYAN M., WIEDLER R., WANSCHITZ I., MENDES H.W., KAYER K.,
SCHUELLER C.;
Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases
[2].
SEQUENCE FROM N.A.
EC ARABIDOPSIS SEQUENCING PROJECT;
Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
EMBL: AL022140; CAAL811.1;
DR MENDEL_29747; Arabid.1088.29747
DR MENDEL_29747; pec_lyase; 1;
FRAN: P00544; pec_lyase; 1;
PRINTS: P00807; AMBALIEPSEN.
LYASE.
KW

Query Match 13.3% Score 135 DB 19 Length 394
Best Local Similarity 36.4% Pred. NO. 1.62e-07

| | Matches | 16: Conservative | 15: Mismatches | 13: Mismatches |
|--|---------|------------------|----------------|----------------|
| DE 105 YGVZCAKPLKWTFAKDWVITLLENLMVNSYKTIIDGRGAKVEZAY | 143 | 13 | 13 | 13 |

| | |
|----|---|
| RE | 9 |
| ID | C65457 PRELIMINARY; PRJ. 394 AA. |
| AC | C65457; |
| DT | D1 01-AUG-1998 (Trevallure) OT, (Gibault) |
| DI | D1 01 AUG 1998 (Trevallure) last sequence obtained |
| DJ | D1 N.V.-1998 (Trevallure) 12/10/98 (Gibault) 04/1998 |
| DE | DATE CLASE LINE PR PROJ |
| EN | FIND-1990. |
| CS | Arabidopsis thaliana (Mouse-ear cress). |
| OC | Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophytes: |
| OC | euphytophytes: Spermatophyta: Magnoliophyta: eudicotyledons: |
| OC | core eudicots: Rosidae: eurosids II: Brassicales: Brassicaceae: |
| OC | Arabidopsids. |
| RN | [1] |
| RP | SEQUENCE FROM N.A. |
| RA | BEVAN M., WEDLER H., WAMBUT E., BANCROFT I., NEWES R.W., MAYER X. |
| PL | SCHUELLER C. |
| FL | Submitted (MAR-1998; to the ENBL/GenBank/CDBJ databases. |
| RN | 12]. |
| RP | SEQUENCE FROM N.A. |
| RA | EU ARABIDOPSIS SEQUENCING PROJECT; |
| RL | Submitted (APR-1998) to the ENBL/GenBank/CDBJ databases. |
| DR | EMBI; AL022140; CAAL812.1; |
| DR | NEMSEL; 29068; Arath;1088;29068. |
| DR | PFAM: PF00544; rev lase; 1 |

PRINIS; PRO0807; AMSALLERGEN.
Lysase.
SEQUENCE 394 AA: 43299 MW; 8B9A93C9 CRC32;

Query Match 13.18; Score 134; DB 10; Length 394;
Best Local Similarity 36.48; Pred. No. 3 58e-07;

Matches 16: Conservative 13: Mismatches 13: Indels 2: Gaps 1:

DB 105 NVVICAKPKITFAKUMVITLANELKVNYSYKTIQSI AKVEIAY 148
 QY 67 PLMIFSGNNKTKRMPMYIAGYKTFDRRAEV 108

RESULT 13
 ID Q23657 PRELIMINARY: PRI: 405 AA.
 AC Q23657
 DT 01-JAN-1998 (FEMBLrel: 05, Created)
 DI 01-JAN-1998 (FEMBLrel: 05, Last sequence update)
 DE 01-NOV-1999 (FEMBLrel: 12, Last annotation update)
 DE PECTATE SPECIFIC PECTATE LYASE (FRAGMENT)
 GN PL.
 OS Arabidopsis thaliana (mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC euphyllophytes; Spermatophyta; Magnoliopsida; eucolecyledons;
 CC core eudicots; Rosidae; eurosids I; Rosales; Rosaceae; Fragaria.
 CC PECTATE LYASE.
 RN 1.
 PP SEQUENCE FROM N.A.
 RP STRAIN: CHANDLER.
 RX MEDLINE: 9435972.
 RA MEDINA ESCOBAR N., CAPERNAS J., MOYANO E., CABALLERO J.L.,
 RA MENZIE-BLANCK J.P.
 RT "Cloning, molecular characterization and expression pattern of a
 RT strawberry ripening-specific cDNA with sequence homology to pectate
 RT lyase from higher plants."
 RI Plant Mol Biol 34: 867-877 (1997).
 DR EMBL: 063551; AAB71208.1; F00544.1
 DR MENDEL: 250544; PECO544; PECO544.1
 DR PRAM: PECO544; pec_lyase: 1
 KW Lyase.
 SQ SEQUENCE 405 AA: 45744 MW: 289944AC CRC32:

Query Match 13.1% Score 134: DB 10: Length 405.
 Best Local Similarity 36.5% Pred No. 3.58e-07:
 Matches 19: Conservative 13: Mismatches 20: Indels 1: Gaps 1:

DB 166 PLMIVKRMKMTITLQELIMNSFKIDARIVNV HAYGSCILIFVIVNI 216
 QY 73 RPKWTFSSNNMKTKRMPYIAGYKTFDRRAEVSYVHNSAKPFRVVGII 124

RESULT 13
 ID Q23657 PRELIMINARY: PRI: 226 AA.
 AC Q23657
 DT 01-JAN-1998 (FEMBLrel: 05, Created)
 DI 01-JAN-1998 (FEMBLrel: 05, Last sequence update)
 DE 01-NOV-1999 (FEMBLrel: 12, Last annotation update)
 DE PUTATIVE PECTATE LYASE (FRAGMENT)
 GN ALC.
 OS Arabidopsis thaliana (mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC euphyllophytes; Spermatophyta; Magnoliopsida; eucolecyledons;
 CC core eudicots; Rosidae; eurosids I; Brassicaceae; Brassicaceae;
 CC Arabidopsis.
 RN 1.
 PP SEQUENCE FROM N.A.
 RP MEDLINE: 97422403.
 RA KULKAUSKAS R., MCCORMACK S.J.
 RT "Identification of the tobacco and Arabidopsis homologues of the
 RT pollen-expressed LAYS9 gene of tomato."
 RI Plant Mol Biol 34: 809-814 (1997).
 DR EMBL: 083521; AAB69761.1; F00544.1
 DR MENDEL: 25561; ARAth1088.25561.
 DR PRAM: PECO544; pec_lyase: 1.
 KW Lyase.
 E1 NON-TER
 F1 NON-TER
 SQ SEQUENCE 226 AA: 25103 MW: 5F1C8DD2 CRC32:

Query Match 13.0% Score 133: DB 10: Length 226:

Best Local Similarity 51.5% Pred No. 5.32e-07:
 Matches 17: Conservative 6: Mismatches 15: Indels 3: Gaps 0:

DB 54 PLMIFGURMTIOLKEELIMNSEKILGKGSV 86
 QY 74 PLMIFSGNNKTKRMPMYIAGYKTFDRRAEV 108

RESULT 12
 ID Q43822 PRELIMINARY: PRI: 438 AA.
 AC Q43822
 DT 01-NOV-1996 (FEMBLrel: 01, Created)
 DI 01-NOV-1996 (FEMBLrel: 01, Last sequence update)
 DE 01-NOV-1999 (FEMBLrel: 12, Last annotation update)
 DE POLLEN SPECIFIC PECTATE LYASE (FRAGMENT)
 DE (PECTATE TRANSELMINASE).
 GN Zea mays (Maize).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC euphyllophytes; Spermatophyta; Magnoliopsida; eucolecyledons;
 CC Poaceae; Zea.
 RN 1.
 PP SEQUENCE FROM N.A.
 RP STRAIN: CV. GOLDEN BUSH.
 RX MEDLINE: 94083558.
 RA TURICH N.P., HAMILLON D.A., MASARENAS J.P.
 RT "Isolation and characterization of pollen-specific cDNA from maize with
 RT sequence homology to rhamnogalacturonate and pectate lyases."
 RI Plant Mol Biol 23: 1061-1065 (1993).
 CC -:- CATALYTIC ACTIVITY: EC:4.1.1.17
 CC GLUCOSACCHARIDES WITH 4-DEXY-ALPHA-D-GLUC-4-EN-1,3-DIOL 4-HS AC
 CC THEIR NON-REDUCING ENDS.
 DR EMBL: L20140; AAA16476.1; F00544.1
 DR MENDEL: 11165; Zeama1059.11165
 DR PRAM: PECO544; pec_lyase: 1
 DR PRINTS: PR00837; AMBALLERGEN.
 KW Lyase.
 SQ SEQUENCE 438 AA: 49148 MW: 1585369 CRC32:

Query Match 12.8% Score 131: DB 10: Length 438
 Best Local Similarity 47.1% Pred No. 1.79e-01:
 Matches 16: Conservative 10: Mismatches 10: Indels 3:

DB 156 RLKIVFARQVIEELKGLIVNNSKIDGRRAEV 189
 QY 73 RPKWTFSSNNMKTKRMPYIAGYKTFDRRAEV 108

RESULT 13
 ID Q43819 PRELIMINARY: PRI: 458 AA.
 AC Q43819
 DT 01-NOV-1996 (FEMBLrel: 01, Created)
 DI 01-NOV-1996 (FEMBLrel: 01, Last sequence update)
 DE PECTATE LYASE (FRAGMENT)
 OS Medicago sativa (Alfalfa).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC euphyllophytes; Spermatophyta; Magnoliopsida; eucolecyledons;
 CC core eudicots; Rosidae; eurosids I; Fabaceae; Papilionaceae;
 CC Medicago.
 RN 1.
 PP SEQUENCE FROM N.A.
 RP MEDLINE: 97422403.
 RA WU Y., GUO X., DU S., ERICKSON L.J.
 RT "Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases
 DR EMBL: U11472; AAB52411.1; F00544.1
 DR MENDEL: 9042; Mcdsa1068.904.
 DR PRAM: PECO544; pec_lyase: 1.
 DR PRINTS: PR00837; AMBALLERGEN.
 KW Lyase.
 SQ SEQUENCE 450 AA: 50155 MW: B39CFD9E CRC32:

Query Match 12.7% Score 130: DB 10: Length 450:
 Best Local Similarity 37.3% Pred No. 1.72e-06:
 Matches 22: Conservative 15: Mismatches 19: Indels 3: Gaps 3:

DB 154 FGAAGCGPGLHIFGSRNVITLIGELVSSDKTIDGRGANYO-IR-DGAG:TMQFYNNV: 216
 QY 67 FTLMGRRPGLHIFSGNNH:KLMKPMYIAGYKTFEGRAEVSVVHVGAKFIR-VGGII 124

RESULT 14
 ID 023668 PRELIMINARY: PRT: 227 AA.
 AC 023668:
 DT 01-JAN-1998 (FEMBLrel: 05, Created)
 DT 01-JAN-1998 (FEMBLrel: 05, Last sequence update)
 DT 01-NOV-1999 (FEMBLrel: 12, Last annotation update)
 DE PUTATIVE PEPTIDE LYASE (FRAGMENT).
 GN All.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC ephylophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 OC core eudicots; Rosidae; euroids 11; Brassicales; Brassicaceae;
 OC Arabidopsis.
 RN [1]
 RF SEQUENCE FROM N.A.
 RX MEDLINE: 97422403.
 RA KULIKAVSKAS R., MCCORMICK S.:
 RI "Identification of the tobacco and Arabidopsis homologues of the
 RI pollen-expressed LARS9 gene of tomato."
 RC Plant Mol. Biol. 34:809-814(1997).
 DR EMBL: U83622; AAF69762.1;
 DR MENDEL: 25562; Arabid1088:25562.
 DE PFAM: PF00544; Pect_lyase: 1.
 KW Lyase.
 FT NON_TER 1 277
 FT NON_TER 227 227
 SQ SEQUENCE 227 AA: 25336 MW: C09A6147 CRC32:

Query Match: 12.6%, Score 129, DB 10: Length 227;
 Best Local Similarity 42.4%, Pred. No. 2.55e-06;
 Matches 14: Conservative 10: Mismatches 9: Indels 0: Gaps 0:

DB 54 PMLVFEKGVIEIKQELIMNSPKTIDARXSNV 85
 QY 74 PMLVFEKGVIEIKQELIMNSPKTIDARXSNV 106

RESULT 5
 ID 024159 PRELIMINARY: PRT: 274 AA.
 AC 024159:
 DT 01-JAN-1998 (FEMBLrel: 05, Created)
 DT 01-JAN-1998 (FEMBLrel: 05, Last sequence update)
 DT 01-NOV-1999 (FEMBLrel: 12, Last annotation update)
 DE PUTATIVE PEPTIDE LYASE N159 (FRAGMENT).
 GN NT59.
 OS Nicotiana tabacum (Common tobacco).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC ephylophytes; Spermatophyta; Magnoliophyta; eudicotyledons;
 OC core eudicots; Asteridae; euasterids 11; Solanales; Solanaceae;
 OC Nicotiana.
 RN [1]
 RF SEQUENCE FROM N.A.
 RX MEDLINE: 97422403.
 RA KULIKAVSKAS R., MCCORMICK S.:
 RI "Identification of the tobacco and Arabidopsis homologues of the
 RI pollen-expressed LARS9 gene of tomato."
 RI Plant Mol. Biol. 34:809-814(1997).
 DR EMBL: U85646; AAB69758.1;
 DR MENDEL: 25645; Nicta:1088:25645.
 DE PFAM: PF00544; Pect_lyase: 1.
 KW Lyase.
 FT NON_TER 274 274
 FT NON_TER 274 274
 SQ SEQUENCE 274 AA: 30790 MW: D2C8C820 CRC32:

Query Match: 12.1%, Score 124, DB 10: Length 274;
 Best Local Similarity 32.1%, Pred. No. 1.75e-05;
 Matches 16: Conservative 16: Mismatches 20: Indels 2: Gaps 2:

DB 154 YAVIQKPELWIIIFERSMTIRLQELIMNSPKTIDARXSNVIAKAGITLQYIKNV 225
 QY 67 FTLMGRRPGLHIFSGNNH:KLMKPMYIAGYKTFEGRAEVSVVHVGAKFIR-VGGII 124

Search completed: Mon Jun 19 15:57:47 2000
 Job time : 29 secs.

WATERMAN

(TM)

Release 1.0A John F. Collins, BioComputing Research Unit,
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

Waterman_HP protein - protein database search, using Smith-Waterman algorithm

Run on: Mon Jun 19 16:02:06 2000 MasPar time 7.21 Seconds
Linear output not generated 439,951 Million cell updates/sec

Title: US-09-142-524A-3
Description: (1226) from US09-42524A.fep
Database: 1 MKVTVAFNGFGRNRVFKRVSNTVTHSRKID:FAKSKPHLQKNTIGTORAKKNNRWLQ 134

Sequence length: 134
Database: 134
Statistics: Mean 30.6127 Variance 134.267 Scale 0.212

Post processing: Minimum Match 0%
Testing first 45 summaries
Maximum DB seq length 134

Database: a-sequences35
sequencesep

Statistics: Mean 30.6127 Variance 134.267 Scale 0.212
Pred. No. is the number of results predicted by charge to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

| Result | No. | Score | Match | Length | DB | ID | Description | Pred. No. |
|--------|------|-------|-------|--------|----|--------|-------------------------|-----------|
| 1 | 1056 | 1056 | 1056 | 134 | 1 | W27371 | Multi-epitope peptide | 4,34e-87 |
| 2 | 541 | 541 | 541 | 134 | 1 | W27370 | Multi-epitope peptide | 6,58e-42 |
| 3 | 372 | 372 | 372 | 134 | 1 | W27369 | Multi-epitope peptide | 5,64e-24 |
| 4 | 288 | 288 | 288 | 134 | 1 | W44682 | T-cell epitope peptide | 5,42e-09 |
| 5 | 188 | 188 | 188 | 134 | 1 | R45545 | Cry 3 I pollen allergen | 1,13e-04 |
| 6 | 158 | 158 | 158 | 134 | 1 | R45579 | Cry 3 I Japanese Cedar | 1,13e-04 |
| 7 | 142 | 142 | 142 | 134 | 1 | R45579 | Cry 3 I pollen allergen | 2,46e-03 |
| 8 | 137 | 137 | 137 | 134 | 1 | R81581 | Cedar pollen allergen | 5,24e-03 |
| 9 | 137 | 137 | 137 | 134 | 1 | W80345 | Sugr allergen protein | 6,33e-03 |

Note: Post processor removed 54 summaries from list due to search parameters chosen.

ALIGNMENTS

RESULT 1
AC W27371 standard, peptide: 134 AA.
CI W27371 1998 (first entry)
DE Multi-epitope peptide used as immunotherapeutic agent #1.
KW Multi-epitope peptide: immunotherapeutic agent: allergic diseases:
KW T-cell epitope region: allergen: lymphocyte: immunoglobulin E.

OS Synthetic.
PN W09732600-A1.
PD 12-SEP-1997.
PF 10-MAR-1997: JP-080702.
PR 10-MAR-1996: JP-080702.
PA (MEIP) MEIJI MILK PROD CO LTD.
PI Dairiki K, Iwama A, Kuno K, Kume A, Some T.
DR WPI: 97-470495/43.
PT Peptide immunotherapeutic agent to treat allergic diseases.
PI Contains multi-epitope peptide containing 1 cell epitope regions
from different allergens.
PS Claim 6: Page 31: 56pp: Japanese.
CC The present sequence represents a multi-epitope peptide which is used as
a new immunotherapeutic agent. It comprises 1 cell epitope regions from 2
or more different allergens (preferably linked via arginine or lysine
dimers), where the 1 cell epitope regions have a positivity index
greater than 100 as measured in a patient group responding to the
allergen: have at least 70% reactivity with lymphocytes from patients
responding to the allergen; and are not reactive with immunoglobulin E
(IgE) antibodies from patients responsive to the allergen. The agent can
be used to prevent and treat a wide variety of allergic diseases, e.g. by
desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
SQ Sequence 134 AA:

Query Match 100.0% Score 1026 DB 1 Length 134
Best Local Similarity 100.0% Pred. NO. 4,34e-87
Matches 134: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
DB 1 MKVTVAFNGFGRNRVFKRVSNTVTHSRKID:FAKSKPHLQKNTIGTORAKKNNRWLQ 50
QY 1 MKVTVAFNGFGRNRVFKRVSNTVTHSRKID:FAKSKPHLQKNTIGTORAKKNNRWLQ 60
DB 61 FAKLTGFTLMGRPELWIFSGUNNINIK:KPMSTAGYKFDGRRRAEVSVYHVNKAF:189V 120
QY 61 FAKLTGFTLMGRPELWIFSGUNNINIK:KPMSTAGYKFDGRRRAEVSVYHVNKAF:189V 120
DB 121 DGIHAYONPASWK 134
QY 121 DGIHAYONPASWK 134

RESULT 2
AC W27370 standard, peptide: 134 AA.
CI W27370 1998 (first entry)
DE Multi-epitope peptide used as immunotherapeutic agent #1.
KW Multi-epitope peptide: immunotherapeutic agent: allergic diseases:
KW T-cell epitope region: allergen: lymphocyte: immunoglobulin E.
OS Synthetic.
PN W09732600-A1.
PD 12-SEP-1997.
PF 10-MAR-1997: 500740.
PR 10-MAR-1996: JP-080702.
PA (MEIP) MEIJI MILK PROD CO LTD.
PI Dairiki K, Iwama A, Kuno K, Kume A, Some T.
DR WPI: 97-470495/43.
PT Peptide immunotherapeutic agent to treat allergic diseases.
PI Contains multi-epitope peptide containing 1 cell epitope regions
from different allergens.
PS Claim 6: Page 31: 56pp: Japan, se.
CC The present sequence represents a multi-epitope peptide which is used as
a new immunotherapeutic agent. It comprises 1 cell epitope regions from 2
or more different allergens (preferably linked via arginine or lysine
dimers), where the 1 cell epitope regions have a positivity index
greater than 100 as measured in a patient group responding to the
allergen: have at least 70% reactivity with lymphocytes from patients
responding to the allergen; and are not reactive with immunoglobulin E
(IgE) antibodies from patients responsive to the allergen. The agent can
be used to prevent and treat a wide variety of allergic diseases, e.g. by
desensitisation. Side effects, e.g. those mediated by IgE, are reduced.
SQ Sequence 105 AA:
Query Match 54.7% Score 561 DB 1 Length 135:

reactive with Japanese cedar pollen allergen. Such modified peptides possess certain characteristics which render them particularly suitable for drug product formulation. Peptide fragments of both modified and unmodified, are given in #8493-#8503. This peptide fragment, crosslinked to amino acids 71-90 of the allergen, was the protein.

Seq. no. 43 AA

2000

Very Much, 5.48; Score 154, 98.1, Length 20

[illegible][illegible]

1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 20. 21. 22. 23. 24. 25. 26. 27. 28. 29. 30. 31. 32. 33. 34. 35. 36. 37. 38. 39. 40. 41. 42. 43. 44. 45. 46. 47. 48. 49. 50. 51. 52. 53. 54. 55. 56. 57. 58. 59. 60. 61. 62. 63. 64. 65. 66. 67. 68. 69. 70. 71. 72. 73. 74. 75. 76. 77. 78. 79. 80. 81. 82. 83. 84. 85. 86. 87. 88. 89. 90. 91. 92. 93. 94. 95. 96. 97. 98. 99. 100. 101. 102. 103. 104. 105. 106. 107. 108. 109. 110. 111. 112. 113. 114. 115. 116. 117. 118. 119. 120. 121. 122. 123. 124. 125. 126. 127. 128. 129. 130. 131. 132. 133. 134. 135. 136. 137. 138. 139. 140. 141. 142. 143. 144. 145. 146. 147. 148. 149. 150. 151. 152. 153. 154. 155. 156. 157. 158. 159. 160. 161. 162. 163. 164. 165. 166. 167. 168. 169. 170. 171. 172. 173. 174. 175. 176. 177. 178. 179. 180. 181. 182. 183. 184. 185. 186. 187. 188. 189. 190. 191. 192. 193. 194. 195. 196. 197. 198. 199. 200. 201. 202. 203. 204. 205. 206. 207. 208. 209. 210. 211. 212. 213. 214. 215. 216. 217. 218. 219. 220. 221. 222. 223. 224. 225. 226. 227. 228. 229. 230. 231. 232. 233. 234. 235. 236. 237. 238. 239. 240. 241. 242. 243. 244. 245. 246. 247. 248. 249. 250. 251. 252. 253. 254. 255. 256. 257. 258. 259. 260. 261. 262. 263. 264. 265. 266. 267. 268. 269. 270. 271. 272. 273. 274. 275. 276. 277. 278. 279. 280. 281. 282. 283. 284. 285. 286. 287. 288. 289. 290. 291. 292. 293. 294. 295. 296. 297. 298. 299. 300. 301. 302. 303. 304. 305. 306. 307. 308. 309. 310. 311. 312. 313. 314. 315. 316. 317. 318. 319. 320. 321. 322. 323. 324. 325. 326. 327. 328. 329. 330. 331. 332. 333. 334. 335. 336. 337. 338. 339. 340. 341. 342. 343. 344. 345. 346. 347. 348. 349. 350. 351. 352. 353. 354. 355. 356. 357. 358. 359. 360. 361. 362. 363. 364. 365. 366. 367. 368. 369. 370. 371. 372. 373. 374. 375. 376. 377. 378. 379. 380. 381. 382. 383. 384. 385. 386. 387. 388. 389. 390. 391. 392. 393. 394. 395. 396. 397. 398. 399. 400. 401. 402. 403. 404. 405. 406. 407. 408. 409. 410. 411. 412. 413. 414. 415. 416. 417. 418. 419. 420. 421. 422. 423. 424. 425. 426. 427. 428. 429. 430. 431. 432. 433. 434. 435. 436. 437. 438. 439. 440. 441. 442. 443. 444. 445. 446. 447. 448. 449. 450. 451. 452. 453. 454. 455. 456. 457. 458. 459. 460. 461. 462. 463. 464. 465. 466. 467. 468. 469. 470. 471. 472. 473. 474. 475. 476. 477. 478. 479. 480. 481. 482. 483. 484. 485. 486. 487. 488. 489. 490. 491. 492. 493. 494. 495. 496. 497. 498. 499. 500. 501. 502. 503. 504. 505. 506. 507. 508. 509. 510. 511. 512. 513. 514. 515. 516. 517. 518. 519. 520. 521. 522. 523. 524. 525. 526. 527. 528. 529. 530. 531. 532. 533. 534. 535. 536. 537. 538. 539. 540. 541. 542. 543. 544. 545. 546. 547. 548. 549. 550. 551. 552. 553. 554. 555. 556. 557. 558. 559. 560. 561. 562. 563. 564. 565. 566. 567. 568. 569. 570. 571. 572. 573. 574. 575. 576. 577. 578. 579. 580. 581. 582. 583. 584. 585. 586. 587. 588. 589. 590. 591. 592. 593. 594. 595. 596. 597. 598. 599. 600. 601. 602. 603. 604. 605. 606. 607. 608. 609. 610. 611. 612. 613. 614. 615. 616. 617. 618. 619. 620. 621. 622. 623. 624. 625. 626. 627. 628. 629. 630. 631. 632. 633. 634. 635. 636. 637. 638. 639. 640. 641. 642. 643. 644. 645. 646. 647. 648. 649. 650. 651. 652. 653. 654. 655. 656. 657. 658. 659. 660. 661. 662. 663. 664. 665. 666. 667. 668. 669. 670. 671. 672. 673. 674. 675. 676. 677. 678. 679. 680. 681. 682. 683. 684. 685. 686. 687. 688. 689. 690. 691. 692. 693. 694. 695. 696. 697. 698. 699. 700. 701. 702. 703. 704. 705. 706. 707. 708. 709. 710. 711. 712. 713. 714. 715. 716. 717. 718. 719. 720. 721. 722. 723. 724. 725. 726. 727. 728. 729. 730. 731. 732. 733. 734. 735. 736. 737. 738. 739. 740. 741. 742. 743. 744. 745. 746. 747. 748. 749. 750. 751. 752. 753. 754. 755. 756. 757. 758. 759. 760. 761. 762. 763. 764. 765. 766. 767. 768. 769. 770. 771. 772. 773. 774. 775. 776. 777. 778. 779. 780. 781. 782. 783. 784. 785. 786. 787. 788. 789. 790. 791. 792. 793. 794. 795. 796. 797. 798. 799. 800. 801. 802. 803. 804. 805. 806. 807. 808. 809. 810. 811. 812. 813. 814. 815. 816. 817. 818. 819. 820. 821. 822. 823. 824. 825. 826. 827. 828. 829. 830. 831. 832. 833. 834. 835. 836. 837. 838. 839. 840. 84

2457, 2458, 2459, 2460, 2461, 2462, 2463, 2464, 2465, 2466, 2467, 2468, 2469, 2470, 2471, 2472, 2473, 2474, 2475, 2476, 2477, 2478, 2479, 2480, 2481, 2482, 2483, 2484, 2485, 2486, 2487, 2488, 2489, 2490, 2491, 2492, 2493, 2494, 2495, 2496, 2497, 2498, 2499, 2500, 2501, 2502, 2503, 2504, 2505, 2506, 2507, 2508, 2509, 2510, 2511, 2512, 2513, 2514, 2515, 2516, 2517, 2518, 2519, 2520, 2521, 2522, 2523, 2524, 2525, 2526, 2527, 2528, 2529, 2530, 2531, 2532, 2533, 2534, 2535, 2536, 2537, 2538, 2539, 2540, 2541, 2542, 2543, 2544, 2545, 2546, 2547, 2548, 2549, 2550, 2551, 2552, 2553, 2554, 2555, 2556, 2557, 2558, 2559, 2560, 2561, 2562, 2563, 2564, 2565, 2566, 2567, 2568, 2569, 2570, 2571, 2572, 2573, 2574, 2575, 2576, 2577, 2578, 2579, 2580, 2581, 2582, 2583, 2584, 2585, 2586, 2587, 2588, 2589, 2590, 2591, 2592, 2593, 2594, 2595, 2596, 2597, 2598, 2599, 2600, 2601, 2602, 2603, 2604, 2605, 2606, 2607, 2608, 2609, 2610, 2611, 2612, 2613, 2614, 2615, 2616, 2617, 2618, 2619, 2620, 2621, 2622, 2623, 2624, 2625, 2626, 2627, 2628, 2629, 2630, 2631, 2632, 2633, 2634, 2635, 2636, 2637, 2638, 2639, 2640, 2641, 2642, 2643, 2644, 2645, 2646, 2647, 2648, 2649, 2650, 2651, 2652, 2653, 2654, 2655, 2656, 2657, 2658, 2659, 2660, 2661, 2662, 2663, 2664, 2665, 2666, 2667, 2668, 2669, 2670, 2671, 2672, 2673, 2674, 2675, 2676, 2677, 2678, 2679, 2680, 2681, 2682, 2683, 2684, 2685, 2686, 2687, 2688, 2689, 2690, 2691, 2692, 2693, 2694, 2695, 2696, 2697, 2698, 2699, 2700, 2701, 2702, 2703, 2704, 2705, 2706, 2707, 2708, 2709, 2710, 2711, 2712, 2713, 2714, 2715, 2716, 2717, 2718, 2719, 2720, 2721, 2722, 2723, 2724, 2725, 2726, 2727, 2728, 2729, 2730, 2731, 2732, 2733, 2734, 2735, 2736, 2737, 2738, 2739, 2740, 2741, 2742, 2743, 2744, 2745, 2746, 2747, 2748, 2749, 2750, 2751, 2752, 2753, 2754, 2755, 2756, 2757, 2758, 2759, 2760, 2761, 2762, 2763, 2764, 2765, 2766, 2767, 2768, 2769, 2770, 2771, 2772, 2773, 2774, 2775, 2776, 2777, 2778, 2779, 2780, 2781, 2782, 2783, 2784, 2785, 2786, 2787, 2788, 2789, 2790, 2791, 2792, 2793, 2794, 2795, 2796, 2797, 2798, 2799, 2800, 2801, 2802, 2803, 2804, 2805, 2806, 2807, 2808, 2809, 2810, 2811, 2812, 2813, 2814, 2815, 2816, 2817, 2818, 2819, 2820, 2821, 2822, 2823, 2824, 2825, 2826, 2827, 2828, 2829, 2830, 2831, 2832, 2833, 2834, 2835, 2836, 2837, 2838, 2839, 2840, 2841, 2842, 2843, 2844, 2845, 2846, 2847, 2848, 2849, 2850, 2851, 2852, 2853, 2854, 2855, 2856, 2857, 2858, 2859, 2860, 2861, 2862, 2863, 2864, 2865, 2866, 2867, 2868, 2869, 2870, 2871, 2872, 2873, 2874, 2875, 2876, 2877, 2878, 2879, 2880, 2881, 2882, 2883, 2884, 2885, 2886, 2887, 2888, 2889, 2890, 2891, 2892, 2893, 2894, 2895, 2896, 2897, 2898, 2899, 2900, 2901, 2902, 2903, 2904, 2905, 2906, 2907, 2908, 2909, 2910, 2911, 2912, 2913, 2914, 2915, 2916, 2917, 2918, 2919, 2920, 2921, 2922, 2923, 2924, 2925, 2926, 2927, 2928, 2929, 2930, 2931, 2932, 2933, 2934, 2935, 2936, 2937, 2938, 2939, 2940, 2941, 2942, 2943, 2944, 2945, 2946, 2947, 2948, 2949, 2950, 2951, 2952, 2953, 2954, 2955, 2956, 2957, 2958, 2959, 2960, 2961, 2962, 2963, 2964, 2965, 2966, 2967, 2968, 2969, 2970, 2971, 2972, 2973, 2974, 2975, 2976, 2977, 2978, 2979, 2980, 2981, 2982, 2983, 2984, 2985, 2986, 2987, 2988, 2989, 2990, 2991, 2992, 2993, 2994, 2995, 2996, 2997, 2998, 2999, 3000, 3001, 3002, 3003, 3004, 3005, 3006, 3007, 3008, 3009, 3010, 3011, 3012, 3013, 3014, 3015, 3016, 3017, 3018, 3019, 3020, 3021, 3022, 3023, 3024, 3025, 3026, 3027, 3028, 3029, 3030, 3031, 3032, 3033, 3034, 3035, 3036, 3037, 3038, 3039, 3040, 3041, 3042, 3043, 3044, 3045, 3046, 3047, 3048, 3049, 3050, 3051, 3052, 3053, 3054, 3055, 3056, 3057, 3058, 3059, 3060, 3061, 3062, 3063, 3064, 3065, 3066, 3067, 3068, 3069, 3070, 3071, 3072, 3073, 3074, 3075, 3076, 3077, 3078, 3079, 3080, 3081, 3082, 3083, 3084, 3085, 3086, 3087, 3088, 3089, 3090, 3091, 3092, 3093, 3094, 3095, 3096, 3097, 3098, 3099, 3100, 3101, 3102, 3103, 3104, 3105, 3106, 3107, 3108, 3109, 3110, 3111, 3112, 3113, 3114, 3115, 3116, 3117, 3118, 3119, 3120, 3121, 3122, 3123, 3124, 3125, 3126, 3127, 3128, 3129, 3130, 3131, 3132, 3133, 3134, 3135, 3136, 3137, 3138, 3139

[illegible]

Japanese cedar: detection: allergy: tree

§ Cryptoceria africana.

20 JAN 1984
15 JAN 1984
10 JAN 1984
01 SEP 1982
01 SEP 1982
(LMD) LMD-107 PEACH CRIP
Rond 15 65mm RZ 31115 11 811 M 11004 11
APR 84 010000Z
01 APR 84

[illegible]

CONFIDENTIAL - ALL INFORMATION CONTAINED HEREIN IS UNCLASSIFIED EXCEPT WHERE SHOWN OTHERWISE

1. The first step is to identify the problem or question that needs to be addressed. This involves understanding the context and the specific requirements of the task.

1. The first step is to identify the problem or question that needs to be addressed. This involves understanding the context and the specific requirements of the task.

2000 1000 500 0

[illegible]

Figure 1: Schematic representation of the experimental design. The diagram shows a sequence of events: 'Stimulus presentation' (a box with a question mark), 'Response' (a box with a question mark), 'Feedback' (a box with a question mark), and 'Inter-trial interval' (a box with a question mark). The sequence is repeated for multiple trials, with a 'Start' box at the beginning and an 'End' box at the end.

.....

1. 2. 3. 4. 5. 6. 7. 8. 9. 10. 11. 12. 13. 14. 15. 16. 17. 18. 19. 20. 21. 22. 23. 24. 25. 26. 27. 28. 29. 30. 31. 32. 33. 34. 35. 36. 37. 38. 39. 40. 41. 42. 43. 44. 45. 46. 47. 48. 49. 50. 51. 52. 53. 54. 55. 56. 57. 58. 59. 60. 61. 62. 63. 64. 65. 66. 67. 68. 69. 70. 71. 72. 73. 74. 75. 76. 77. 78. 79. 80. 81. 82. 83. 84. 85. 86. 87. 88. 89. 90. 91. 92. 93. 94. 95. 96. 97. 98. 99. 100. 101. 102. 103. 104. 105. 106. 107. 108. 109. 110. 111. 112. 113. 114. 115. 116. 117. 118. 119. 120. 121. 122. 123. 124. 125. 126. 127. 128. 129. 130. 131. 132. 133. 134. 135. 136. 137. 138. 139. 140. 141. 142. 143. 144. 145. 146. 147. 148. 149. 150. 151. 152. 153. 154. 155. 156. 157. 158. 159. 160. 161. 162. 163. 164. 165. 166. 167. 168. 169. 170. 171. 172. 173. 174. 175. 176. 177. 178. 179. 180. 181. 182. 183. 184. 185. 186. 187. 188. 189. 190. 191. 192. 193. 194. 195. 196. 197. 198. 199. 200. 201. 202. 203. 204. 205. 206. 207. 208. 209. 210. 211. 212. 213. 214. 215. 216. 217. 218. 219. 220. 221. 222. 223. 224. 225. 226. 227. 228. 229. 230. 231. 232. 233. 234. 235. 236. 237. 238. 239. 240. 241. 242. 243. 244. 245. 246. 247. 248. 249. 250. 251. 252. 253. 254. 255. 256. 257. 258. 259. 260. 261. 262. 263. 264. 265. 266. 267. 268. 269. 270. 271. 272. 273. 274. 275. 276. 277. 278. 279. 280. 281. 282. 283. 284. 285. 286. 287. 288. 289. 290. 291. 292. 293. 294. 295. 296. 297. 298. 299. 300. 301. 302. 303. 304. 305. 306. 307. 308. 309. 310. 311. 312. 313. 314. 315. 316. 317. 318. 319. 320. 321. 322. 323. 324. 325. 326. 327. 328. 329. 330. 331. 332. 333. 334. 335. 336. 337. 338. 339. 340. 341. 342. 343. 344. 345. 346. 347. 348. 349. 350. 351. 352. 353. 354. 355. 356. 357. 358. 359. 360. 361. 362. 363. 364. 365. 366. 367. 368. 369. 370. 371. 372. 373. 374. 375. 376. 377. 378. 379. 380. 381. 382. 383. 384. 385. 386. 387. 388. 389. 390. 391. 392. 393. 394. 395. 396. 397. 398. 399. 400. 401. 402. 403. 404. 405. 406. 407. 408. 409. 410. 411. 412. 413. 414. 415. 416. 417. 418. 419. 420. 421. 422. 423. 424. 425. 426. 427. 428. 429. 430. 431. 432. 433. 434. 435. 436. 437. 438. 439. 440. 441. 442. 443. 444. 445. 446. 447. 448. 449. 450. 451. 452. 453. 454. 455. 456. 457. 458. 459. 460. 461. 462. 463. 464. 465. 466. 467. 468. 469. 470. 471. 472. 473. 474. 475. 476. 477. 478. 479. 480. 481. 482. 483. 484. 485. 486. 487. 488. 489. 490. 491. 492. 493. 494. 495. 496. 497. 498. 499. 500. 501. 502. 503. 504. 505. 506. 507. 508. 509. 510. 511. 512. 513. 514. 515. 516. 517. 518. 519. 520. 521. 522. 523. 524. 525. 526. 527. 528. 529. 530. 531. 532. 533. 534. 535. 536. 537. 538. 539. 540. 541. 542. 543. 544. 545. 546. 547. 548. 549. 550. 551. 552. 553. 554. 555. 556. 557. 558. 559. 560. 561. 562. 563. 564. 565. 566. 567. 568. 569. 570. 571. 572. 573. 574. 575. 576. 577. 578. 579. 580. 581. 582. 583. 584. 585. 586. 587. 588. 589. 590. 591. 592. 593. 594. 595. 596. 597. 598. 599. 600. 601. 602. 603. 604. 605. 606. 607. 608. 609. 610. 611. 612. 613. 614. 615. 616. 617. 618. 619. 620. 621. 622. 623. 624. 625. 626. 627. 628. 629. 630. 631. 632. 633. 634. 635. 636. 637. 638. 639. 640. 641. 642. 643. 644. 645. 646. 647. 648. 649. 650. 651. 652. 653. 654. 655. 656. 657. 658. 659. 660. 661. 662. 663. 664. 665. 666. 667. 668. 669. 670. 671. 672. 673. 674. 675. 676. 677. 678. 679. 680. 681. 682. 683. 684. 685. 686. 687. 688. 689. 690. 691. 692. 693. 694. 695. 696. 697. 698. 699. 700. 701. 702. 703. 704. 705. 706. 707. 708. 709. 710. 711. 712. 713. 714. 715. 716. 717. 718. 719. 720. 721. 722. 723. 724. 725. 726. 727. 728. 729. 730. 731. 732. 733. 734. 735. 736. 737. 738. 739. 740. 741. 742. 743. 744. 745. 746. 747. 748. 749. 750. 751. 752. 753. 754. 755. 756. 757. 758. 759. 760. 761. 762. 763. 764. 765. 766. 767. 768. 769. 770. 771. 772. 773. 774. 775. 776. 777. 778. 779. 780. 781. 782. 783. 784. 785. 786. 787. 788. 789. 790. 791. 792. 793. 794. 795. 796. 797. 798. 799. 800. 801. 802. 803. 804. 805. 806. 807. 808. 809. 810. 811. 812. 813. 814. 815. 816. 817. 818. 819. 820. 821. 822. 823. 824. 825. 826. 827. 828. 829. 830. 831. 832. 833. 834. 835. 836. 837. 838. 839. 840. 84

08-03-19

cedar pollen allergen peptide

| Case | Age | Sex | Site of tumor | Antibody | Pathologic therapy | Immunotherapy |
|------|-----|-----|---------------|----------|--------------------|---------------|
| 1 | 33 | F | Colon | CEA | None | None |
| 2 | 40 | M | Colon | CEA | None | None |
| 3 | 40 | M | Colon | CEA | None | None |
| 4 | 40 | M | Colon | CEA | None | None |
| 5 | 40 | M | Colon | CEA | None | None |
| 6 | 40 | M | Colon | CEA | None | None |
| 7 | 40 | M | Colon | CEA | None | None |
| 8 | 40 | M | Colon | CEA | None | None |
| 9 | 40 | M | Colon | CEA | None | None |
| 10 | 40 | M | Colon | CEA | None | None |
| 11 | 40 | M | Colon | CEA | None | None |
| 12 | 40 | M | Colon | CEA | None | None |
| 13 | 40 | M | Colon | CEA | None | None |
| 14 | 40 | M | Colon | CEA | None | None |
| 15 | 40 | M | Colon | CEA | None | None |
| 16 | 40 | M | Colon | CEA | None | None |
| 17 | 40 | M | Colon | CEA | None | None |
| 18 | 40 | M | Colon | CEA | None | None |
| 19 | 40 | M | Colon | CEA | None | None |
| 20 | 40 | M | Colon | CEA | None | None |
| 21 | 40 | M | Colon | CEA | None | None |
| 22 | 40 | M | Colon | CEA | None | None |
| 23 | 40 | M | Colon | CEA | None | None |
| 24 | 40 | M | Colon | CEA | None | None |
| 25 | 40 | M | Colon | CEA | None | None |
| 26 | 40 | M | Colon | CEA | None | None |
| 27 | 40 | M | Colon | CEA | None | None |
| 28 | 40 | M | Colon | CEA | None | None |
| 29 | 40 | M | Colon | CEA | None | None |
| 30 | 40 | M | Colon | CEA | None | None |
| 31 | 40 | M | Colon | CEA | None | None |
| 32 | 40 | M | Colon | CEA | None | None |
| 33 | 40 | M | Colon | CEA | None | None |
| 34 | 40 | M | Colon | CEA | None | None |
| 35 | 40 | M | Colon | CEA | None | None |
| 36 | 40 | M | Colon | CEA | None | None |
| 37 | 40 | M | Colon | CEA | None | None |
| 38 | 40 | M | Colon | CEA | None | None |
| 39 | 40 | M | Colon | CEA | None | None |
| 40 | 40 | M | Colon | CEA | None | None |
| 41 | 40 | M | Colon | CEA | None | None |
| 42 | 40 | M | Colon | CEA | None | None |
| 43 | 40 | M | Colon | CEA | None | None |
| 44 | 40 | M | Colon | CEA | None | None |
| 45 | 40 | M | Colon | CEA | None | None |
| 46 | 40 | M | Colon | CEA | None | None |
| 47 | 40 | M | Colon | CEA | None | None |
| 48 | 40 | M | Colon | CEA | None | None |
| 49 | 40 | M | Colon | CEA | None | None |
| 50 | 40 | M | Colon | CEA | None | None |
| 51 | 40 | M | Colon | CEA | None | None |
| 52 | 40 | M | Colon | CEA | None | None |
| 53 | 40 | M | Colon | CEA | None | None |
| 54 | 40 | M | Colon | CEA | None | None |
| 55 | 40 | M | Colon | CEA | None | None |
| 56 | 40 | M | Colon | CEA | None | None |
| 57 | 40 | M | Colon | CEA | None | None |
| 58 | 40 | M | Colon | CEA | None | None |
| 59 | 40 | M | Colon | CEA | None | None |
| 60 | 40 | M | Colon | CEA | None | None |
| 61 | 40 | M | Colon | CEA | None | None |
| 62 | 40 | M | Colon | CEA | None | None |
| 63 | 40 | M | Colon | CEA | None | None |
| 64 | 40 | M | Colon | CEA | None | None |
| 65 | 40 | M | Colon | CEA | None | None |
| 66 | 40 | M | Colon | CEA | None | None |
| 67 | 40 | M | Colon | CEA | None | None |
| 68 | 40 | M | Colon | CEA | None | None |
| 69 | 40 | M | Colon | CEA | None | None |
| 70 | 40 | M | | | | |

14-511-1995: J2-2

(HAYB) HAYASHI BARA SEIBOTSU KAGAKU.

5/6/60 4:25 PM

allergen-variant

WQSRQL (TM)

Release: 1A John F. Collins, PhD, Patina Research Unit,
Copyright (c) 1994-1998 University of Edinburgh, UK.
Distribution rights by Oxford Molecular Ltd

Message: Protein - protein database search using Smith-Waterman algorithm
Run on: Tue Jun 20 13:38:51 2000 Master time 4.74 seconds
408.006 Million cell updates/sec
Larger output not generated.

File: 05000142524A.p
Protein: 111341 (100) USC9142524A.p
Residues: 1226
Sequence: MKVTVAENFGNRRVFIRK.....KFIKRVLSIIAAYQNPASR 134
Scoring table: BAW 150
Gap 11

Searches: 145341 seqs, 14437400 residues

Post-processed: Minimum Ratio: 0%
Listing first 45 summaries
Maximum DB seq length: 134

Database: released
111341 MB 215B COMB 4167 COMB 51backfiles

Statistics: Mean 9.72% Variance 129.31% scale 0.222

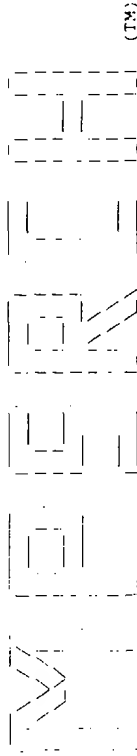
Proj. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARY

| Result | Copy | No. Score | Match length | DB ID | Description | Proj. No. |
|--------|------|-----------|--------------|-------|-------------|-----------|
|--------|------|-----------|--------------|-------|-------------|-----------|

No matches found.

Search completed: Tue Jun 20 13:38:57 2000
Job time 1.6 secs.



(TM)

Release 3.1A John F. Collins, BioComputing Research Unit,
Copyright (c) 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MPearchApp protein - protein database search, using Smith-Waterman algorithm
Run On: Mon Jun 19 16:02:39 2000: MasPar time 30.42 Seconds
443.817 Million cell updates/sec
Linear output not generated.

Index: US-09-142-524A-3
Description: (1194) from US09142524A.rap
Perfect Score: 1026
Sequence: 1 MKVTVAFNFGPNRVRVEIKR.....KFI5RVGGITAAAYONPASWK 134

Scoring table: RAW 1500
Gap 11

Searched: 721238 seqs, 100765575 residues

Post-processing: Minimum Match: 0%
Listing first 45 summaries
Maximum DB seq length 134

Database:

appending
1:09:27 2106 31060 4107 51080 61081 71082 81083 91084
10:084B 11:085 12:086 13:087 14:088 15:089 16:090 17:091
18:092 19:093 20:094 21:095 22:NEWP 23:NEWJ60 24:NEWJ8
25:NEWJ9

Statistics: Mean 13.909: Variance 128.683: scale 0.254

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query % | Match | Length | DB ID | Description | Pred. No. |
|------------|---------|-------|--------|-------|-----------------------------------|-----------|
| 1 | 1026 | 100.0 | 134 | 17 | US-09-142- Sequence 3, Applicatio | 1.61e-97 |
| 2 | 561 | 54.7 | 105 | 17 | US-09-142- Sequence 2, Applicatio | 2.24e-46 |
| 3 | 472 | 26.3 | 80 | 17 | US-09-142- Sequence 1, Applicatio | 3.77e-26 |
| 4 | 296 | 28.8 | 60 | 9 | US-08-350- Sequence 62, Applicati | 3.14e-18 |
| 5 | 296 | 28.8 | 60 | 7 | US-08-236- Sequence 62, Applicati | 3.14e-18 |
| 6 | 296 | 28.8 | 60 | 10 | US-08-467- Sequence 62, Applicati | 3.14e-18 |
| 7 | 296 | 28.8 | 60 | 10 | US-08-467- Sequence 62, Applicati | 3.14e-18 |
| 8 | 296 | 28.8 | 60 | 4 | US-07-938- Sequence 62, Applicati | 3.14e-18 |
| 9 | 296 | 28.8 | 60 | 10 | US-08-468- Sequence 62, Applicati | 3.14e-18 |
| 10 | 296 | 28.8 | 60 | 10 | US-08-467- Sequence 62, Applicati | 3.14e-18 |
| 11 | 172 | 16.9 | 128 | 8 | US-08-350- Sequence 187, Applicat | 7.10e-06 |
| 12 | 172 | 16.8 | 128 | 10 | US-08-468- Sequence 187, Applicat | 7.10e-06 |
| 13 | 172 | 16.8 | 128 | 10 | US-08-467- Sequence 187, Applicat | 7.10e-06 |
| 14 | 172 | 16.8 | 128 | 10 | US-08-467- Sequence 187, Applicat | 7.10e-06 |

Note: Post-processor removed all summaries from list due to search parameters chosen.

ALIGNMENTS

RESULT 1
ID US-09-142-524-3 STANDARD: PRT: 134 AA.
XX xxxxxx
AC
XX
XX
DT
XX
XX
DE
XX
XX

Sequence 3, Application US/09142524

Sequence 3, Application US/09142524

GENERAL INFORMATION:

APPLICANT: Sone, Toshio

APPLICANT: Kure, Akimori

APPLICANT: Kairiki, Kazuo

APPLICANT: Iwama, Akiko

APPLICANT: Kihno, Kohsuke

TITLE OF INVENTION: Peptide-based Immunotherapeutic Agent For Treating

TITLE OF INVENTION: Allergic Diseases

FILE REFERENCE: Docket No. SPO-103

CURRENT APPLICATION NUMBER: US/09/142.524

CURRENT FILING DATE: 1999-1-04

EARLIER APPLICATION NUMBER: 8/60/702

EARLIER FILING DATE: 1996-03-10

EARLIER APPLICATION NUMBER: FCT/JP97/00740

EARLIER FILING DATE: 1997-03-10

NUMBER OF SEQ ID NOS: 5

SOFTWARE: Patentin Ver. .

SEQ ID NO 3

LENGTH: 134

TYPE: PRT

ORGANISM: Unknown

FEATURE:

OTHER INFORMATION: Description of Unknown Organism peptide

SEQUENCE 134 AA: 15546 MW: 90017 CN:

Query Match 100.0% Score 1026: DB 17: Length 134

Best Local Similarity 100.0% Pred. No. 1.61e-97

Matches 134: Conservative 0: Mismatches 0: Indels 0: Gaps 0

DB 1 MKVTVAFNFGPNRVRVEIKRVSNIHSRDIIFASKNFHLKNTIGGPKKSNPKLQ 10

QY 1 MKVTVAFNFGPNRVRVEIKRVSNIHSRDIIFASKNFHLKNTIGGPKKSNPKLQ 10

DB 61 FAKLQGTIMGRPLKLTFSNNKTKKEPMYIAKYKTFERRAEVSVVNNKATIRPV 120

QY 61 FAKLQGTIMGRPLKLTFSNNKTKKEPMYIAKYKTFERRAEVSVVNNKATIRPV 120

DB 61 FAKLQGTIMGRPLKLTFSNNKTKKEPMYIAKYKTFERRAEVSVVNNKATIRPV 120

QY 61 FAKLQGTIMGRPLKLTFSNNKTKKEPMYIAKYKTFERRAEVSVVNNKATIRPV 120

DB 121 DGIIAAYONPASWK 134

QY 121 DGIIAAYONPASWK 134

RESULT 2
ID US-09-142-524-2 STANDARD: PRT: 105 AA.
XX xxxxxx
AC
XX
XX
DT
XX
XX
DE
XX
XX

Sequence 2, Application US/09142524

Sequence 2, Application US/09142524

GENERAL INFORMATION:

APPLICANT: Sone, Toshio

APPLICANT: Kure, Akimori

APPLICANT: Kairiki, Kazuo

APPLICANT: Iwama, Akiko

APPLICANT: Kihno, Kohsuke

TITLE OF INVENTION: Peptide-based Immunotherapeutic Agent For Treating

TITLE OF INVENTION: Allergic Diseases

FILE REFERENCE: Docket No. SPO-103

CURRENT APPLICATION NUMBER: US/09/142.524


```
##model_type=DNA
##residues=1731 ##label=RES
##cross-references EMBL:U10529; NID:Q511451; PIR:Q511451
SUMMARY      #length 132 #molecular_weight 15592 #checksum 9163

Query Match:      8 94 Score 91 DB 2: Length 131
Best Local Similarity 25.34 Pred No. 5,456-00
Matches 160 Conservative 211 Mismatches 301 Indels 31 Gaps 31

DB      10 DUFFINERANDWITSNVYVLLGQWYVNTYSSKAKLKAALGSGKG-KKSRILRRF 68
        1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
CY      4 TVAPNGGPNRAVE-IRKVSNVTHGRRIQ-IFASKNFLOKNTISIGRWKNNRIWQGF 61
        1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
IR      69 TRAGYVILAIR 80
        1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QJ      62 AKLTQFLNGRR 73
```

Search completed: Mon Jun 19 16:01:59 2000
Run time 1.15 secs.

WIRELESS

(TM)

Release 4.1A John F. Collins, BioComputing Research Unit,
Copyright 1993-1998 University of Edinburgh, U.K.
Distribution rights by Oxford Molecular Ltd

MSearch_PP protein - protein database search, using Smith-Waterman algorithm
Run on: Mon Jun 19 16:00:24 2000: Maspar time 7.92 Seconds
Tabular output not generated. 515 241 Million cell updates/sec

Title: US-09-142-524A-3
Description: (-134) from US09.42524A.pef
Perfect Score: 1245
Sequence: 1 MKVTVAFNQSPNRRVFIR.....KEIFRVEDGIIAAYQNPASWK 134

Scoring table: IAM 150
Gap 11

Searched: 8887 seqs, 30454973 residues
Post-processing: Minimum Match 0%
Listing first 45 summaries
Maximum DA seq length 134

Database: SWISS-prot:38
Travisprot

Statistics: Mean 44.072: Variance 79.368: scale 0.555

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

| SUMMARIES | | | | |
|------------|-------------|--------------|-------|-----------------------------------|
| Result No. | Query Score | Match Length | DA ID | Description |
| 1 | 105 | 10 2 | 110 1 | Y12K_SMSV4 POSSIBLE 12 KD NUCLEIC |

Note: Post-processor removed 44 summaries from list due to search parameters chosen.

ALIGNMENTS

RESULT 1

| | | | | |
|----|--|-----------------------------------|------|---------|
| ID | Y12K_SMSV4 | STANDARD: | PR1: | 110 AA. |
| AC | P36289: | | | |
| DT | 01-JUN-1994 | (Rel. 29, Created) | | |
| DI | 01-JUN-1994 | (Rel. 29, Last sequence update) | | |
| D1 | 15-DEC-1998 | (Rel. 37, Last annotation update) | | |
| DE | POSSIBLE 12 KD NUCLEIC ACID-BINDING PROTEIN. | | | |
| DS | San Miguel sea lion virus (serotype 4) (SMSV 4). | | | |
| OC | Viruses; ssRNA positive-strand viruses, no DNA stage: Caliciviridae; | | | |
| CC | Calicivirus. | | | |
| LN | 11 | | | |
| RN | SEQUENCE FROM N.A. | | | |
| RF | MEDLINE: 92410750. | | | |
| RA | Neill J.D.: | | | |
| RT | Nucleotide sequence of the capsid protein gene of two serotypes of | | | |

RT San Miguel sea lion virus: identification of conserved and conserved amino acid sequences among calicivirus capsid proteins
RL Virus Res. 24:211-222(1992)
CC -!- SIMILARITY: TO FELINE CALICIVIRUS 12 KD PROTEIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Data Bank, the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is not in any way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isit.isit.ac.uk/usage.htm> or send an email to license@isit.ac.uk)
CC
CC EMBL: M67482; AAA16221.1;
DR PIR: D48562; S48562.
DR
SQ SEQUENCE 110 AA: 12566 MW: 14255559482741EE 09464

Query Match: 10.2% Score 105 DA 11 Length 110
Best Local Similarity 31.9% Prod No: 159e-02
Matches 15: Conservative 15: Mismatches 14: Indels 4: Gaps 2:

Db 10 FANSVANAVVEGKKDL-ASKGLQKNSRALDTERAFNYHM--IAFEK 53
CY 17 FKKKSNVILHGRRIIDIFASKNFHCKNTIGTGRKXKNPWLQFAK 64

Search completed: Mon Jun 19 16:00:34 2000
Job time : 10 secs.

 WATSON

Released via John F. Collins, BioComputing Research Unit,
 Copyright (c) 1993-1998 University of Edinburgh, U.K.
 Distribution Rights by Oxford Molecular Ltd

Miscblif protein - protein database search, using Smith-Waterman algorithm
 Run on: Mon Jun 19 16:00:52 2000; Maspar time 18.49 Seconds
 502.347 Million cells updates/sec
 Individual output not generated.

Title: US-09-142-524A-3
 Description: (1114); from US09142524A.pep
 Perfect Score: 1426
 Sequence: MKVIVAFNPGPRNRVFIKR.....KPIRRVGGIIAAYQNPASKK 134

Scoring matrix: BAW 150
 Gap 12

Searched: 125876 seqs, 69334122 residues

Fast-processing: Minimum Match 0%
 Listing first 45 summaries
 Maximum DB seq length 134

Database:

41 items in 2
 1. sp_archaea 2.sp_bacteria 3.sp_fungi 4.sp_human
 5.sp_invertebrate 6.sp_mammal 7.sp_phc 8.sp_ornithine
 9.sp_plant 10.sp_plant 11.sp_rodent 12.sp_unclassified
 13.sp_vertebrate 14.sp_virus

Statistics: Mean 42.980; Variance 75.878; scale 0.566

Pred No. is the number of results predicted by chance to have a
 score greater than or equal to the score of the result being printed,
 and is derived by analysis of the total score distribution.

| Result | Query | Match | Length | ID | Description | Pred. No. |
|--------|-------|-------|--------|----|------------------------|-----------|
| 1 | 95 | 93 | 108 | 14 | CAPSID | 5.52e-01 |
| 2 | 90 | 89 | 55 | 1 | PROTEIN TRANSLOCASE, S | 2.76e-00 |

Note: Post-processor removed 43 summaries from list due to search parameters chosen.

ALIGNMENTS

RESULT 1
 ID C09783 PRELIMINARY: PRT: 108 AA.
 AC C09783
 DI 01-JUL-1997 (TrEMBLrel 04, Created)
 DI 01-JUL-1997 (TrEMBLrel 04, Last sequence update)
 DI 01-NOV-1998 (TrEMBLrel 08, Last annotation update)
 DE CAPSID.
 OS Saccharomyces cerevisiae virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
 CC Calicivirus.
 FN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=SMSV-17;
 RX MEDLINE: 98187296;
 RA WATSON D.O., BERKE T., GINSBURG M.B., POST S.E., ZHENG W.X., LAM K.M.,
 RA JIANG X., GOLDING B., SMITH A.W.,
 RT "Partial characterization of the genome of nine calici-
 RI caliciviruses".
 RL Arch. Virol. 141:2443-2456(1996);
 CR EMBL: U52005; AAC57041.1;
 SQ SEQUENCE 108 AA: 12354 MW: 476345.5 GPC12

Query Match 9.8% Score 90; DB 14; Length 108
 Best Local Similarity 34.0% Pct. No. 5.52e-01
 Matches 14; Conservative 14; Mismatches 15; Indels 1; Gaps 2;

DB 10 FUNGVANAAIECKKIDLASG LKNAALGKMKENYAR-LAFER SA
 QY 17 EIKRVSNVIMHRRIDIPASG LQANTGUTERRWKNRWIQFAK SA

RESULT 2
 ID C29714 PRELIMINARY: PRT: 55 AA.
 AC C29714

DI 01-JAN-1996 (TrEMBLrel 05, Created)
 DI 01-JAN-1996 (TrEMBLrel 05, Last sequence update)
 DI 01-AUG-1998 (TrEMBLrel 07, Last annotation update)
 DE PROTEIN TRANSLOCASE, SUBUNIT SEC61 GAMMA (SET3).
 GN AF0536.

OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
 CC Archaeoglobus.

RP SEQUENCE FROM N.A.
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49612;
 RX MEDLINE: 98049343

RA KLECK H.P., CLAYTON R.A., JUNG J.-P., WHITE G., NELSON K.B.,
 RA KETCHUM K.A., DOOSON R.J., OWEN M., HICKY E.K., PETERS N.C.,
 RA RICHARDSON D.L., KERAVAGE A.R., GRAHAM D.E., KYLESEN N.C.,
 RA FLEUSCHMANN R.D., QUACKENBUSH J., LEE H.H., SUTHERLAND J.,
 RA KIRKNESS E.F., DOUGHERTY J., WICKENY K., ADAMS M.D., LITTON S.

RA PETERSON S., REICH C.J., MONTEILLO K., BAUMER J.H., GOSLER A., REED L.,
 RA OVERBEER K., GOCAYNE J.D., WEIDMAN J.F., McDONALD L., OTTERBACH T.,
 RA COTTON M.D., SPRIGGS T., ALLACH P., KAINE B.P., SYKES S.M.,
 RA SADOW P.W., D'ANDREA K.P., ROMAN C., FUJII C., GARLAND S.A.,
 RA WASTEN T.M., OLSEN G.J., F. SER C.M., SMITH R.W., WESLEY C.R.,
 RA VENTER J.C.

RC "The complete genome sequence of the hyperthermophilic, extremophile
 P1 refeeding archaeon Archaeoglobus fulgidus".
 PL Nature 392:164-170(1997).

CR EMBL: A001067; AAB4734.1;
 CR TIGR: AF0536.

KW Hypothetical protein.

SQ SEQUENCE 55 AA: 7547 MW: 93.887EA GRC12.

Query Match 8.8% Score 90; DB 14; Length 66;
 Best Local Similarity 36.7% Pct. No. 2.76e-00;
 Matches 11; Conservative 7; Mismatches 11; Indels 1; Gaps 1;

DB 24 MAKKPDWEEFSMTAKVALAV-NEIVGFVGF 52

QY 70 MORRPDUIFSSNNMKMKMPYIAGYKIF 99

Search completed: Mon Jun 19 16:01:18 2000
 Job time: 26 secs.